ISTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

MODIFIED HILBERT HUANG TRANSFORM FOR DATA ANALYSIS AND ITS APPLICATION TO VIBRATION SIGNALS

Ph.D. Thesis by
Ahmet ÖZTÜRK

Department : Electrical Engineering

Programme : Electrical Engineering

MARCH 2011

ISTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

MODIFIED HILBERT HUANG TRANSFORM FOR DATA ANALYSIS AND ITS APPLICATION TO VIBRATION SIGNALS

Ph.D. Thesis by
Ahmet ÖZTÜRK
(504022008)

Date of submission : **15 June 2010**
Date of defence examination: **28 March 2011**

Supervisor (Chairman) : **Prof. Dr. Serhat ŞEKER (ITU)**
Members of the Examining Committee : **Prof. Dr. Bilge GÜNSEL (ITU)**
Prof. Dr. A. Faik MERGEN (ITU)
Prof. Dr. A. Hamdi KAYRAN (ITU)
Assoc. Prof. Dr. Onur TOKER (FU)

MARCH 2011

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

DEĞİŞTİRİLMİŞ HİLBERT HUANG DÖNÜŞÜMÜ İLE VERİ ANALİZİ ve TİTREŞİM SİNYALLERİNE UYGULAMASI

DOKTORA TEZİ

Ahmet ÖZTÜRK
(504022008)

Tezin Enstitüye Verildiği Tarih : 15 Haziran 2010
Tezin Savunulduğu Tarih : 28 Mart 2011

Tez Danışmanı :	**Prof. Dr. Serhat ŞEKER (İTÜ)**
Diğer Jüri Üyeleri :	**Prof. Dr. Bilge GÜNSEL (İTÜ)**
	Prof. Dr. A. Faik MERGEN (İTÜ)
	Prof. Dr. A. Hamdi KAYRAN (İTÜ)
	Doç. Dr. Onur TOKER (FÜ)

MART 2011

FOREWORD

Because of the probable but unintentional simplification in the following thanks giving, I would like to present my thanks and apologies for anyone, in advance. It is for sure that nothing would be at this particular point in time and space if any of the premises would differ.

First of all, I would like to express my deepest appreciations and thanks for my advisor Prof. Serhat Şeker, Prof. Ali Demir and Prof. Mehmet Karaca whose support and patience made this study reach conclusion. I also would like to thank to the members of the examining committee, Prof. Bilge Günsel and Prof. Ahmet Faik Mergen, for their invaluable advises and time spent to support this study.

I would like to thank Prof. B.R. Upadhyaya and his research team at the University of Tennessee, Nuclear Engineering Department, for permission to use the experimental data concerned in this study.

I also would like to present my thanks to NASA and Prof. Huang for their leading and inspiring studies and particularly for support in providing the “Hilbert Huang Transform Data Processing System” software and its documentation.

I am grateful to my family for their constant support, thoughtfulness and belief in me.

Finally, I would like to present my thanks to my university, Istanbul Technical University and its Institute of Science and Technology, that I feel honour to be a part of.

Certains disent que "si belle"...
Que s'il ya un doute ou quelque chose de plus belle.
Donc, je dis “oui belle” et "la seule beauté" sans aucun doute.
La beauté divine et a promis de ma vie...

March 2011 Ahmet ÖZTÜRK

TABLE OF CONTENTS

Page

ABBREVIATIONS

3D	**:** Three Dimensional
CDF	**:** Characteristic Defect Frequency
CMF	**:** Combined Mode Function
CWT	**:** Continuous Wavelet Transform
DFT	**:** Discrete Fourier Transform
EMD	**:** Empirical Mode Decomposition
FFT	**:** Fast Fourier Transform
HHT	**:** Hilbert Huang Transform
HT	**:** Hilbert Transform
IEEE	**:** The Institute of Electrical and Electronics Engineers
IMF	**:** Intrinsic Mode Function
IO	**:** Index of Orthogonality
mHHT	**:** modified Hilbert Huang Transform
NASA	**:** National Aeronautics and Space Administration
PSD	**:** Power Spectral Density
RMS	**:** Root Mean Square
SD	**:** Standard Deviation
STFT	**:** Short Time Fourier Transform
WT	**:** Wavelet Transform

LIST OF TABLES

LIST OF FIGURES

Page

MODIFIED HILBERT HUANG TRANSFORM FOR DATA ANALYSIS AND ITS APPLICATION TO VIBRATION SIGNALS

SUMMARY

The Hilbert Huang Transformation, which is an empirical method currently and known as an efficient method in analysis of non-stationary and non-linear signals due to its data-driven nature, constitutes the basis for this thesis. Within the thesis:

- Attributed fundamental concepts of data analysis are briefly mentioned,
- The original Hilbert Huang Transformation is presented with attribution to various studies,
- A new algorithm, which concerns modifications in the formulation of the local mean approximation and adjustment of boundary values, is proposed,
- Frequency resolution capabilities of the original and modified transformations are presented comparatively over analysis of a deterministic signal,
- Vibration signals acquired from an electric motor, which is subject to an experimental accelerated aging process, are analyzed with:
    - The Fourier Transform,
    - The Continuous Wavelet Transform,
    - The Hilbert Huang Transform, and
    - The Modified Hilbert Huang Transform.

The first and general justification to include the Continuous Wavelet Transformation in this study is, its theoretical completeness. The second and particular one is, the original and published results obtained in analyses of the mentioned vibration signals with the Continuous Wavelet Transform.

The major impacts of the proposed modifications in the algorithm are increased resolution in identification of intrinsic frequencies and reduction in number of iterations to identify them, respectively.

Planned future work is stated with justifications in the conclusion part of the thesis.

DEĞİŞTİRİLMİŞ HİLBERT HUANG DÖNÜŞÜMÜ İLE VERİ ANALİZİ ve TİTREŞİM SİNYALLERİNE UYGULAMASI

ÖZET

Veri uyarlamalı doğası sebebiyle doğrusal ve durağan olmayan sinyallerin analizinde etkinliği ile bilinen ve halihazırda tamamen deneysel bir yöntem olan Hilbert Huang Dönüşümü, tez çalışmasının temelini oluşturmaktadır. Tez çalışmasında:

- İlgili temel veri analiz kavramlarına kısaca değinilmiş,
- Çeşitli çalışmalara atıf ile özgün Hilbert Huang Dönüşümü tanıtılmış,
- Dönüşümün yerel ortalama yaklaşımının ve sınır değerlerinin değiştirilmesine dönük yeni bir algoritma önerilmiş,
- Deterministik bir sinyalin analizi üzerinden özgün ve değiştirilmiş dönüşümlerin frekans çözünürlüğü karşılaştırmalı olarak sunulmuş,
- Deneysel hızlandırılmış yaşlandırma sürecindeki bir elektrik motorundan alınan titreşim işaretleri,
    - Fourier Dönüşümü
    - Sürekli Dalgacık Dönüşümü
    - Hilbert Huang Dönüşümü ve
    - Değiştirilmiş Hilbert Huang Dönüşümü ile analiz edilmiştir.

Tez çalışmasında Sürekli Dalgacık Dönüşümü'ne yer verilmesinin ilk ve genel gerekçesi, teorisinin tamamen belirli olmasıdır. İkinci ve özel gerekçe ise bahsi geçen titreşim işaretlerinin Sürekli Dalgacık Dönüşümü ile analizlerinde elde edilen ve yayınlanan özgün sonuçlarıdır.

Algoritmada önerilen değişikliklerin ana etkileri, içsel kip tanımlamada artan frekans çözünürlüğü ve yineleme sayısındaki anlamlı azalmadır.

İleriye dönük çalışmalar, tezin sonuç bölümünde gerekçeli olarak belirtilmiştir.

1. INTRODUCTION

The Hilbert Huang Transform is a technique that had been developed by Prof. Huang to address time frequency analysis of nonlinear and non-stationary data and appreciated as one of the most important discoveries by NASA in the field of applied mathematics in its history [1].

The technique, consisting of two phases, first exploits the "Empirical Mode Decomposition" (EMD) which is supposed to yield a finite number of "Intrinsic Mode Functions" (IMF) constituting the components of the analyzed signal. In the second phase, Hilbert Transforms of the IMFs are calculated. Well derived IMFs that satisfy the definitive conditions can be expected to admit well behaved Hilbert Transforms. The result is designated as the Hilbert spectrum and presents an energy-frequency-time distribution that does not include any mathematically generated spurious content as the FFT does [2].

With the innovative IMF definition, the data is let to provide the basis for the decomposition. The release of requirement for a priori basis function definition [3] makes the technique adaptive and suitable for analysis of non-stationary signals.

In addition, the technique breaks through the limitation of the uncertainty principle inherited in the Fourier Transform pairs, or Fourier type of transform pairs such as the Wavelet transform [4].

With regards to the briefly stated features above, the technique has attracted enormous interest and has been applied to various areas such as vibration signals, financial time series, physiological data, geophysical data, climatologic data, oceanographic data, etc.

The effectiveness and success of the technique is stated in the literature in comparison with the methods such as FFT and Wavelet Transform. Nevertheless, the technique still includes outstanding problems for improvement.

1.1 Purpose of the Thesis

The thesis is a forward step taken, for the scholar group led by Prof. Şeker, in the analysis of vibration data acquired from an electric motor subject to an aging process. For the specifications of the process [5, 6] can be referred. In the previous studies of the group, the data was subject to various analyses whose results were presented as articles and theses. The important aspects of the results obtained with Wavelet Transform during the studies are presented in the related section to establish the reference for benchmarking.

Development of thorough understanding and software of HHT technique for the analysis of the vibration data is the first of the two main objectives of this study, whereas, the presentation of improvement proposals concerning the HHT technique and obtained results is the second.

1.2 Background

As the systems get more complicated and expensive with demand for high reliability and precision, the significance of condition monitoring, feature extraction and fault diagnosis techniques increases, making the signal analysis indispensable. The subject signals and the signal processing methods vary. Nevertheless, machine vibration data is one of the most widely used signal for condition monitoring [7]. The conducted studies estimates that more than 50% of the failures are mechanical in nature, such as bearing, balance and alignment related problems [8]. The analyses of vibration signals also showed that the state and condition of systems can be identified in association with characteristic frequencies. As an example, the bearings, which are the biggest single cause of motor failures [8, 9], generate vibrations of particular frequencies as they get degraded or even they are fault-free [8, 10]. And the processing methods employed in the analyses can be categorized with respect to the domain of the representation such as time domain, frequency domain, and time-frequency domain. Articles including general assessments of the methods as well as particular applications of time domain methods such as Root Mean Square (RMS), Skewness, Kurtosis, Crest Factor and frequency domain methods such as Fourier Transform, Power Cepstrum, Denoising can be found in the literature. In a review, which had considered the condition monitoring of rolling element bearings using vibration analysis, the frequency domain parameters were found to be more consistent in the detection of damage than time domain parameters [11]. However, with sufficient evidence, it was concluded that it would be unreliable to depend exclusively on any one technique to detect bearing damage.

Joint analysis of time and frequency to identify the energy distribution constitutes another alternative in the toolbox of signal analysis. The Short Time Fourier Transform (STFT), Wigner-Ville Distribution and Wavelet Transforms are some of the time-frequency analysis methods. Of those methods, Wavelet Transform is the most widely used one for bearing fault diagnosis [8, 12].

1.3 Hypothesis

EMD method of HHT can produce IMFs with better frequency resolution and this can be achieved by modifying the local mean approximation and boundary values. Performance of the method can be verified by synthetic and actual vibration data.

2. DATA ANALYSIS

A brief knowledge on the attributes of data analysis techniques and their properties are given as follows.

2.1 Objectives

In its evolutionary track to be synthesized as knowledge, data represent the glimpse of reality acquired including some imperfections within. Exclusion or mitigation of the imperfections and extraction of information are the inevitable parts of data analysis. Put on the agenda by the Chaos Theory, the "Butterfly Effect" explains the demolishing effect of negligence and emphasizes the necessity of precision in the analysis. With the motivation to be precise and to be in compliance with reality, scientific research concerns the invention of better methods and practical applications, hence validation, of the proposed methods. In accordance with the perspective stated above, data analysis serves two purposes [2]:

- Determination of the parameters needed to construct the necessary model,
- Validation of the constructed model to represent the phenomenon.

Prominent source of data are the non-stationarity, non-linearity, noise effects and data span. And in association, the difficulties concerning the methods are adaptivity and locality. Also, resolution in both time and frequency appears as an inevitable requirement. In the following sections, these topics will be dealt briefly.

2.2 Stationarity

Stationarity is a concept that implies whether the statistical descriptors, the moments, of a time series varies with respect to time or not. It is a general requirement for most of the available data analysis methods. According to the traditional definition, a time series, $x(t)$, is stationary in the wide sense, if, for all *t*, **(2.1)-(2.3)** are satisfied.

$$E(|x(t)^2|) < \infty \quad \textbf{(2.1)}$$

$$E\big(x(t)\big) = m \tag{2.2}$$

$$C\big(x(t_1), x(t_2)\big) = C\big(x(t_1 + \tau), x(t_2 + \tau)\big) = C(t_1 - t_2) \tag{2.3}$$

In which $E(\cdot)$ is the expected value defined as the ensemble average of the quantity, and $C(\cdot)$ is the covariance function. Stationarity in the wide sense is also known as weak stationarity or second order stationarity. A time series, $x(t)$, is strictly stationary, if the joint distribution of $[x(t_1), x(t_2), \ldots, x(t_n)]$ and $[x(t_1 + \tau), x(t_2 + \tau), \ldots, x(t_n + \tau)]$ are the same for all t_i and τ. Thus, a strictly stationary process with finite second moments is also weakly stationary, but the inverse is not true [2].

As a result of finite data span, these values need to be estimated. The first four moments of a time series are the mean, covariance, skewness and kurtosis, respectively. For a given data set $\{x_i\}$ the estimates of the moments are defined by **(2.4)-(2.7)** and these statistical parameters maybe used to perform a quick check of the changes in the statistical behavior of a signal [12]:

1st moment:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{2.4}$$

2nd moment:

$$\sigma = \sqrt{\frac{1}{N}\sum\nolimits_{i=1}^{N} (x_i - \mu)^2} \tag{2.5}$$

3rd moment:

$$c = \frac{\left[\frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^3\right]}{\sigma^3} \tag{2.6}$$

4th moment:

$$k = \frac{\left[\frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^4\right]}{\sigma^4} \tag{2.7}$$

Besides the others, the Kurtosis can be used as a standalone bearing condition indicator. Because, the vibration data of a healthy bearing exhibit a normal distribution whereas damage of the bearings cause peakedness in the probability distribution of the data [10, 13].

2.3 Linearity

An operator A is said to be linear if, for any set of admissible values $x_1, x_2, x_3, \dots, x_N$ and constants $a_1, a_2, a_3, \dots, a_n$, **(2.8)** is satisfied [14]:

$$A\left[\sum_{i=1}^{N} a_i x_i\right] = \sum_{i=1}^{N} a_i A[x_i] \quad \textbf{(2.8)}$$

In words, the operation is both additive and homogeneous.

2.4 Locality and Adaptivity

The necessary conditions for the basis to represent a nonlinear and non-stationary time series are summarized as [2]:

- Completeness: Guarantee of the degree of precision of the expansion.
- Orthogonality: Guarantee of positivity of energy and avoidance of leakage.
- Locality: Provides identification of all events by the time of their occurences.
- Adaptivity: Local variations of the data are regarded and adapted to maintain the decomposition fully account for the underlying physics of the processes and not just to fulfil the mathematical requirements for fitting the data.

Locality and adaptivity are the necessary conditions for the basis for expanding nonlinear and non-stationary time series; The principle of this basis construction is based on the physical time scales that characterize the oscillations of the phenomena.

2.5 Fourier Transform

For a signal or function $f(t)$, the Fourier Transform is defined by **(2.9)** and the inverse Fourier Transform to recover the signal or function is defined by **(2.10)**.

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt \tag{2.9}$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{j\omega t}d\omega \tag{2.10}$$

These two Fourier Transform equations are called a Fourier Transform pair. The original function is in the time domain, *t*, which is transformed into the frequency domain, *ω*. The resulting transformed signal gives an indication of the frequency components that contribute to the original signal when a spectral plot of amplitude versus frequency is constructed. The function *f(t)* can be continuous or discrete; however, the preceding definition for the Fourier Transform intrinsically assumes *f(t)* to be continuous. The Discrete Fourier Transform (DFT) of a discrete signal, *f(t)*, designated as f_k, is defined by **(2.11)**

$$F_n = \frac{1}{N}\sum_{k=1}^{N} f_k e^{{-2\pi jnk}/{N}} \tag{2.11}$$

And the inverse Fourier Transform by **(2.12)**.

$$f_k = \sum_{n=0}^{N-1} F_n e^{{2\pi jnk}/{N}} \tag{2.12}$$

where *n* and *j* are indices [15].

2.6 Wavelet Transform and Signal Decomposition

The wavelet transform is a tool that splits data into different frequency components, and then each component can be studied with a resolution matched to its scale. The use of the wavelet transform is efficient for fault diagnosis, since the technique gives the information about the signal in the time and the frequency domains.

The crucial aspect in wavelet analysis is the selection of the basis function, in other words the mother wavelet. The dependence of a proper time-frequency-energy distribution to basis selection is one of the two major disadvantages of Wavelet Transform. The second disadvantage is that resolution in both time and frequency cannot be provided.

The Continuous Wavelet Transform (CWT) of a signal $x(t)$ is defined by **(2.13)**

$$CWT(\alpha,\tau) = \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{+\infty} x(t)\Psi_{\alpha,\tau}^{*}(t)dt \qquad \textbf{(2.13)}$$

where $\Psi^{*}(t)$ denotes the complex conjugate of mother wavelet $\Psi(t)$ such that

$$\Psi_{\alpha,\tau}(t) = \frac{1}{\sqrt{\alpha}}\Psi\left(\frac{t-\tau}{\alpha}\right) \qquad \textbf{(2.14)}$$

Parameters α and τ are scale and translation indices which corresponds to mutual frequency and the time shifting, respectively.

2.7 Wavelet Packet Transform

The use of Wavelet Packets for Event Detection WPT is an extension of Discrete Wavelet Transform and can be obtained by a generalization of the fast pyramidal algorithm. Each detail coefficient vector is decomposed into two parts using the same approach as in approximation vector splitting. The complete binary tree is produced as shown in Figure 2.1.

We start with *h(n)* and *g(n)*, the two impulsive responses of low-pass and high-pass analysis filters, corresponding to the scaling function and the wavelet function, respectively. The sequence of functions is defined by **(2.15)** and **(2.16)** [16]:

$$W_{2n}(x) = \sqrt{2}\sum_{k=0}^{2N-1} h(k)W_n(2x-k) \qquad \textbf{(2.15)}$$

$$W_{2n+1}(x) = \sqrt{2}\sum_{k=0}^{2N-1} g(k)W_n(2x-k) \qquad \textbf{(2.16)}$$

where $W_0(x) = \phi(x)$ is the scaling function and $W_1(x) = \Psi(x)$ is the wavelet function. In other words, the three indexed family of analyzing functions can be reached by **(2.17)**.

$$W_{j,n,k}(x) = 2^{-j/2} W_n(2^{-j}x - k) \qquad \textbf{(2.17)}$$

k can be interpreted as a time localization parameter and j as a scale parameter. $W_{j,n,k}$ analyzes the fluctuations of the signal roughly around the position $2^j k$, at the scale 2^j, and at various frequencies for the different admissible values of the last parameter n[1]. The wavelet packets' coefficients at each node (j, n) are given by **(2.18)**.

$$\left(C_{jn}(k)\right) = \langle f(t), 2^{-j/2} W_n(2^{-j}t - k) \rangle \qquad \textbf{(2.18)}$$

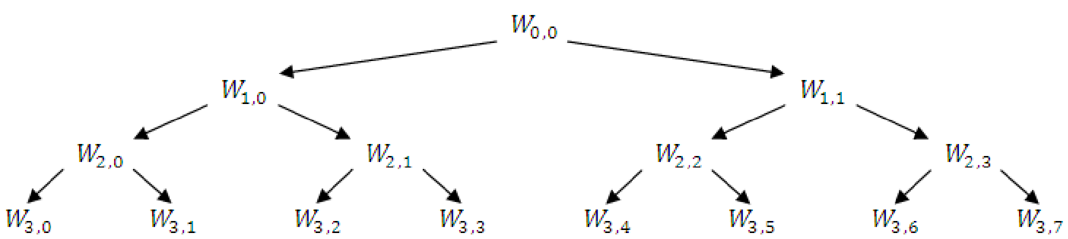

Figure 2.1: Wavelet Packet Decomposition Tree.

3. HILBERT HUANG TRANSFORM

In this chapter, the Hilbert Huang Transform is introduced as an alternative method to the previously mentioned analysis methods such as Fourier and Wavelet Transforms.

3.1 Objectives

Presentation of the listed subjects below constitute the objective of this chapter:

- HHT method and its conceptual components,
- Comparative analysis with FFT and WT,
- A brief survey concerning the essential features, open problems, some related improvement proposals,
- A brief review of current applications and related assessments.

3.2 Conceptual Components of the Method

As it will be presented throughout this chapter, despite its deficiencies, analysis of a signal with HHT is a relatively simple but effective method in comparison to its widely used and analytically well developed alternatives FFT and WT.

The principle of exploiting a signal with complex oscillations to reveal its intrinsic modes empirically is the first of the two phases of the method. Realization of this pure principle is provided in a simple iterative process of EMD. The decomposition is followed by the time-frequency localization of the oscillations which is the second phase of the method and achieved by the Hilbert Transform.

These two conceptual components will be dealt in the following subsections with an emphasis on the EMD since the Hilbert Transform is a mathematically definite transform with the unique property of creating an analytic function as the resultant. But, in precedence of the two, the innovative concept introduced by the method and the natural product of the EMD, the Intrinsic Mode Function is defined.

3.2.1 Intrinsic mode function

A time series, collection of data samples, generally include oscillations particular to the underlying dynamics of a system which result in interlaced local extrema. The oscillations may be around either a constant or varying local mean with a charasteristic time scale. The alternatives for defining a time scale may be:

- The time between successive zero-crossings [2,17];
- The time between successive extrema [2,17];
- The time between successive curvature extrema [2,17].

The transform adopts the time lapse between successive extrema as the characteristic time scale and the two definitive conditions for IMF are given as [2, 18, 19]:

i. The number of extrema and the number of zero crossings must either be equal or differ at most by one in the entire data set.
ii. At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The implications of these conditions are symmetry in wave profiles and local zero mean, respectively. As the conditions are satisfied within EMD process, an oscillatory signal defining the basis function, particular for the analyzed signal, will be achieved and definition of a meaningful instantaneous frequency will be enabled.

3.2.2 Empirical mode decomposition

The motivation behind the EMD approach is that the signal is comprised of coexisting oscillatory modes and can be decomposed into IMFs as a result of a sifting process. An analogy can be established between the terms of the sifting processes of Wavelet and Hilbert Huang Transforms. In Wavelet Transform, at each level of decomposition, the filtering process separates high and low frequency components into detail and approximation terms, respectively, and the sifting process continues over the approximation term. In a similar manner, the sifting process in HHT/EMD behaves like a high pass filter and removes the low frequency components iteratively until a predefined criteria met to yield the IMF component that contain high frequencies. The removal of the IMF component from the analyzed signal yields the residual term over which the sifting process continues. Therefore, in a manner, it can be concluded that the approximation and detail terms of WT correspond to the residual and IMF component terms, respectively. Despite the similarity in principle, the computation and separation of high and low frequency contents through the sifting operation essentially differ as it is explained in the following paragraphs.

The first step of the sifting process is the identification of the local extrema of the signal. The local mean is approximated by computing the local average of the two cubic splines whose knots are the local maxima and minima, respectively. The upper and lower envelopes are required to cover all the data between them. The difference between the signal and the local mean approximation yields the new signal to be sifted until a termination criterion is met.

The first and second steps of the process, with the same rationale applicable to the n^{th} step, is illustrated in Figure 3.1 and 3.2, respectively, using a signal model $x(t)$ comprised by superimposing two sine waves and a bias term as per **(3.1)**.

$$x(t) = \sin(2\pi f_1 t) + \sin(2\pi f_2 t) + 2 \tag{3.1}$$

The data is depicted in Figure 3.1 in blue color. The upper and the lower envelopes are depicted in Figure 3.1 in green and red colors, respectively. The local mean vector obtained as the local average of the envelopes is depicted in Figure 3.1 in black. The signal obtained as the difference between the data and the local mean is depicted in the last section of Figure 3.1 in blue color since this signal constitutes the data that the sifting process continues over as it is illustrated in Figure 3.2.

Reiteration of sifting with the high pass data of the previous step defines a much more compliant new signal with the definitive conditions of IMF. As it can be verified from steps depicted in Figure 3.2, the local mean vector approaches zero and the oscillations around it are almost symmetric. Eventually, as the sifting process is iterated k times, an IMF shall be identified. With removal of the identified IMF from the original signal, the residual data is obtained. The residual data constitutes the new input to the sifting process to yield another IMF.

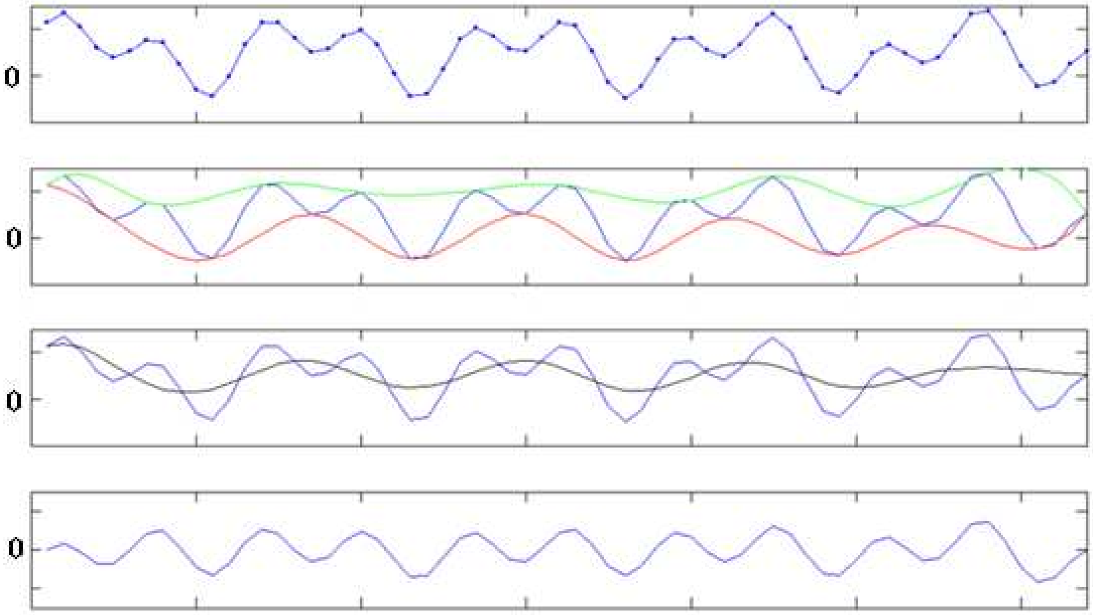

Figure 3.1: The sifting process – 1st iteration.

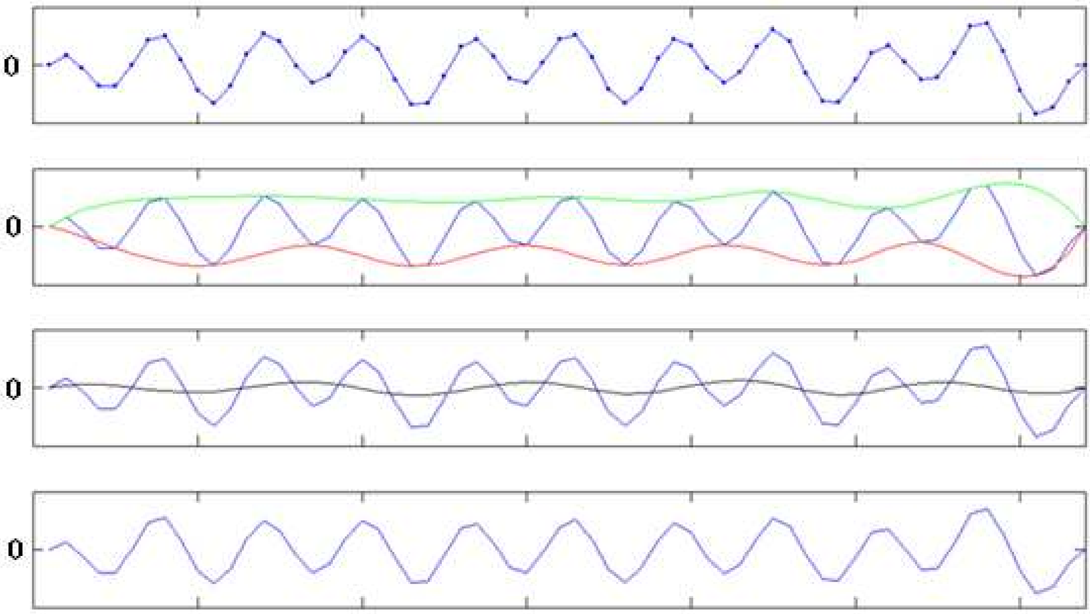

Figure 3.2: The sifting process – 2nd iteration.

The sifting process is intrinsically defined for continuous signals. But the practical applications concern discrete data. Thus, the time index in the computations needs to be handled with care. The algorithm is as follows:

(0) The data is retained in vector x,

(1) Identification of the local extrema,

(2) Generation of the envelopes:

   a. A cubic spline is fit to local maxima to generate the upper envelope e_u,

   b. e_u is resampled at original data locations,

   c. A cubic spline is fit to local minima to generate the lower envelope e_l,

   d. e_l is resampled at original data locations,

(3) Computation of local mean m_1 at corresponding data locations:

$$m_1 = \frac{e_u + e_l}{2} \tag{3.2}$$

(4) Computation of the first IMF candidate:

The difference between the m_1 and the analyzed signal gives the first IMF candidate.

$$h_1 = x - m_1 \tag{3.3}$$

(5) Check for the termination criteria:

Ideally, h_1 should be an IMF [2]. If so, h_1 is assigned as the first IMF c_1,

$$c_1 = h_1 \tag{3.4}$$

Otherwise, h_1 is taken as the analyzed signal and the steps (1)-(3) are reiterated *k* times for $h_1$until termination criteria is satisfied. After satisfaction of the termination criteria, h_{1k} is defined as the first IMF.

$$h_{1k} = h_{1(k-1)} - m_{1k} \tag{3.5}$$

$$c_1 = h_{1k} \tag{3.6}$$

The termination decision is given with regards to *SD* value calculated by

$$SD = \sum_{t=0}^{T} \left[\frac{\left| h_{1(k-1)} - h_{1k} \right|^2}{h_{1(k-1)}^2} \right] \tag{3.7}$$

Typical values to stop the process are proposed to be 0.2-0.3. The nominator term and hence *SD* value will decrease as the sifting process eliminate the riding waves and smooths the uneven amplitudes [2].

(6) Removal of IMF and computation of the residual data:

As the first IMF is defined, it is removed from the original signal to obtain the residual term r_1.

$$r_1 = x - c_1 \tag{3.8}$$

Now, the residual data constitutes the new input to the sifting process. Steps (1)-(4) are repeated to obtain c_2. The removal of the newly identified component c_2. shall yield the new residual term r_2:

$$r_2 = r_1 - c_2 \tag{3.9}$$

In generalized terms,

$$r_i = r_{i-1} - c_i \text{ where } i = 1,2, \dots, n \tag{3.10}$$

In equation (3.8), x is used only in the first IMF computation, then it is replaced with r_i.

(7) *n* IMFs shall be obtained until the decomposition process is stopped. The criterion for stopping may either be reaching a prescribed number of IMFs or r_n becomes a monotonic function or a constant that is not suitable for sifting. The transform is complete and allows reconstruction of the original signal from the components plus the n^{th} residual term r_n as per **(3.11)**:

$$x = r_n + \sum_{i=1}^{n} c_i \tag{3.11}$$

Consequently, the EMD algorithm has two loops in which residual terms and IMFs are computed. Since the inner loop that runs for IMF identification is executed indefinitely (step (1)-(5)) until the termination criteria is satisfied, it creates the most of the computational cost. The calculation of the residual in the outer loop is just a subtraction operation (step (6)) which is done once per each identified IMF.

3.2.3 Hilbert Transform

In 1743, Leonard Euler derived the formula given in **(3.12)**.

$$e^{jz} = cos(z) + j\, sin(z) \tag{3.12}$$

150 years later, the complex notation of harmonic wave forms are formulized by Arthur E: Kennelly and Charles P. Steinmetz as in **(3.13)**:

$$e^{j\omega t} = cos(\omega t) + j\, sin(\omega t) \tag{3.13}$$

Later on, in the beginning of the 20th century, David Hilbert showed that the function $sin(\omega t)$ is the Hilbert Transform of $cos(\omega t)$. This gave us the $\pm {}^{\pi}/_{2}$ phase-shift operator which is a basic property of the Hilbert Transform [20].

The Hilbert Transform $Y(t)$ of a function $x(t)$ is defined for all t in **(3.14)**.

$$Y(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \tag{3.14}$$

when the integral exists. P indicates the Cauchy Principal Value. This transform exists for all functions of class L^p. A real function $x(t)$ and its Hilbert Transform $Y(t)$ are related to each other in such a way that they together create a so called strong analytic signal and they are orthogonal. For a strong analytic signal *f(x),* it is given by **(3.15)** that [20]:

$$H\{Re[f(x)]\} = Im[f(x)] \tag{3.15}$$

Meaning that, $Y(t)$ and $x(t)$ form the complex conjugate pair. Hence, an analytic signal $Z(t)$ can be defined as **(3.16)** [2, 21, 22],

$$Z(t) = x(t) + iY(t) \tag{3.16}$$

The strong analytic signal can be written with an amplitude and a phase

$$Z(t) = a(t)e^{i\theta(t)} \tag{3.17}$$

in which

$$a(t) = [x^2(t) + Y^2(t)]^{1/2}, \tag{3.18}$$

$$\theta(t) = arctan\left(\frac{Y(t)}{x(t)}\right) \quad \textbf{(3.19)}$$

and by equation **(3.20)**, the derivative of the phase $\theta(t)$ can be identified as the instantaneous frequency ω.

$$\omega = \frac{d\theta(t)}{dt} \quad \textbf{(3.20)}$$

Even though a function and its Hilbert Transform are orthogonal, it is not always realized in applications because of truncations in numerical calculations.

3.3 Application Areas of the Method

The HHT method has a wide spectrum of application areas within the scope of signal processing. Some of the topics include [3]:

- Noise filtering,
- Vibration, speech, and acoustic signal analyses,
- Machine health monitoring,
- Non-destructive evaluation for structural health monitoring,
- Financial market data analysis,
- Earthquake engineering, seismic data,
- Bio-medical applications,
- Satellite data,
- Cosmological gravity wave and planets hunting.

3.4 Features and Outstanding Problems

The general methodological outstanding problems of signal and system analysis concern adaptive data analysis, nonlinear system identification and prediction problem for nonstationary processes [3].

The features of HHT, due to which the method is appreciated as an innovative and a promising method for nonlinear and nonstationary data will be dealt in this section. It should be reminded that these features may also be the address of outstanding problems and thus, subjects of improvements.

3.4.1 Orthogonality, uniqueness and independence

Orthogonality requires "inner product", which is a generalization of the notion of the dot product between two complex vectors, be equal to zero.

$$\langle x, y \rangle = 0 \qquad \textbf{(3.21)}$$

By virtue of the decomposition, the elements should all be locally orthogonal to each other. The two elements for checking can be taken as [2]:

i. The local mean vector $\overline{x(t)}$ which was corresponding to m of the , of the signal calculated as the local average of the upper and lower envelopes, and

ii. The difference vector obtained by removing m from the signal.

In this circumstance [2],

$$\overline{\left(x(t) - \overline{x(t)}\right) \cdot \overline{x(t)}} = 0 \qquad \textbf{(3.22)}$$

Hence, it can be concluded that the IMFs are theoretically orthogonal. Since local mean is computed through the envelopes, it may deviate from the true mean. Hence, Equation (3.22) is not strictly true and leakage is unavoidable. However, leakage should be small. In reality, even pure sinusoidal components with different frequencies are not exactly orthogonal due to the finite data length [2].

Huang had defined a measure for orthogonality check which is called "Index of Orthogonality" (IO) and proposed two methods for numerical calculation of it [2]. The first of the methods was for overall and the second was for any two components such as C_f and C_g. They are defined as follows [2]:

$$IO = \sum_{t=0}^{T}\left(\sum_{j=1}^{n+1}\sum_{k=1}^{n+1} C_j(t)C_k(t)/x^2(t)\right) \quad \textbf{(3.23)}$$

$$IO_{fg} = \sum_{t}\frac{C_f C_g}{C_f^2 + C_g^2} \quad \textbf{(3.24)}$$

It should be noted that, although the definition of orthogonality seems to be global, the real meaning here applies only locally [2] implying that no two IMFs will have the same frequency at the same point in time. The local orthogonality property of the decomposition helps to maintain the independence of basis functions. The sum of two IMFs will usually not yield another IMF except under the circumstance that the two IMF components are close in frequency [18].

3.4.2 Frequency resolution

Frequency (spectral) resolution concerns the capability of distinguishing between neighbouring spectral components. Although there are numerical differences, the various envelope processes are shown to yield approximately similar IMF components so that without attention to the choice of function actually employed to construct the signal envelope, the procedure is claimed to give unique frequency resolution [18].

In his studies concerning the spectral resolution of EMD process, Rilling analyzed separability of two or more tones. Three of the conclusions included in the study are as follows [23]:

- The amplitude ratio affects the separation. When the amplitude of the high frequency component gets smaller than the amplitude of the low frequency component, the performance drastically decreases.
- Above the critical cut-off frequency ratio ($f_c = 0.67$) it is impossible to separate the two components, whatever the amplitude ratio is.
- The performance of the method is improved in the presence of a third component.

For further details, [23] can be referred. In this study, the conclusion concerning the frequency ratio is illustrated in Section 4.5 and used for the purpose of validation of the algorithms.

3.4.3 Mode mixing

One of the major drawbacks of the original EMD is the frequent appearance of mode mixing, which is defined as a single Intrinsic Mode Function (IMF) either consisting of signals of widely disparate scales, or a signal of a similar scale residing in different IMF components [24]. Hence, relation of mode mixing with signal intermittency, that can be encountered as a consequence of arbitrary and abrupt occurrences or inoccurences of a component in the time history of a signal, is obvious. The EMD is not able to separate this phenomenon [10, 24]. Also, straightforward implementation of the sifting procedure is stated to produce mode mixing that results in aliasing in the IMFs [3].

Some of the proposed methods to overwhelm this problem include application of intermittence test [3], preprocessing with wavelets[25], a masking approach [10] and a noise aided data analysis approach [24].

3.4.4 Boundary effects (End effects)

The inclusion of the first and the last terms of the signal as both minima and maxima and thus as knots for the cubic spline to be fitted, initiates swings on both tips of the signal that eventually propagate inward on each iteration and corrupt the data. The phenomenon is called the end effects or boundary effects [2, 26]. With respect to the propagation depth, this effect may deteriorate the Hilbert Transform at least or totally corrupts the data.

Concerning studies and proposed improvements generally include addition of boundary extrema in various ways. Some of the proposed methods, with the general intention to represent characteristic natural behaviors of the original time series [27], are as follows:

- Additional waves based on the two closest maxima and minima [2, 27],
- Padding with additional ‘‘characteristic’’ or ‘‘typical’’ waves based on the closest maximum and minimum [28].

$$wave\ extension = A.sin\left(\frac{2\pi t}{p} + \ phase\right) + \ local\ mean \qquad \textbf{(3.25)}$$

where the typical amplitude, A, and period, P, are determined by the nearest local extrema.

- Additional waves are obtained by mirror symmetry of the extrema closest to the edge [27, 29].
- Slope-based method - Generation of new extrema are based on the slopes of extrema near both ends of the data series [27].
- Improved slope-based method (ISBM) [27].

3.4.5 The sifting

With reference to the arguments of the sifting process, which are the local mean (m_{ik}) and the difference (h_{ik}), Huang addresses two problems:

i. Ideally, h_1 should be an IMF. The method for the construction of h_1 described above seems to have been made to satisfy all the requirements of IMF. In reality, however, overshoots and undershoots are common, which can also generate new extrema, and shift or exaggerate the existing ones [2].

ii. Still another complication is that the envelope mean may be different from the true local mean for nonlinear data; consequently, some asymmetric wave forms can still exist no matter how many times the data are sifted [2].

3.4.6 The intrinsic mode functions

The frequency resolution capability of EMD was mentioned previously in Section 3.4.2. Sometimes, the IMFs may not be mono-component functions due to various reasons, such as:

- As the EMD cannot decompose narrowband multi-harmonic signal very well, most of the IMFs are still multi-harmonic functions [30].
- The EMD may generate some undesired low amplitude IMFs at the low-frequency region and raise some undesired frequency components [31].
- The first IMF may cover a wide frequency range at the high-frequency region [31].

- The EMD operation often cannot separate some low-energy components from the analyzed signal, therefore those components may not be able to appear in the frequency–time plane [31].

One more deficiency to be mentioned within this section concerns over-decomposition of a signal. It has been shown that EMD may split an intrinsic mode into two or more neighbouring IMFs because of the effect of noises so that some intrinsic mode functions are distorted and fail to denote the signal features [19].

The proposed solutions include either pre-processing or post-processing of the acquired or derived signals, such as:

- Wavelet Packet Transform, which is a sophisticated tool on its own already as it is briefly explained in Section 2.7, can be given as a successful example for preprocessing. At first step, the original signal is preprocessed with the wavelet packet transform to obtain a set of narrow band signals. Then, the narrow band signals are processed with EMD. As the last step prior to Hilbert Transformation, the set of IMF candidates needs to be filtered for selection of appropriate IMFs. A detailed implementation is given in [31] concluding that a better resolution both in time domain and in frequency domain had been achieved.

- Combined Mode Function (CMF) approach, an example for post processing, is shown as an effective way to represent the true features of the mode obtained by combining over-decomposed components [19].

3.4.7 Effect of sampling rate

Identification of the extrema in the signal is the primary step in EMD. Lacking extrema cause wrong definition of the characteristic time scale and consequently, wrong calculation of both the IMF and the local mean. Therefore, the adverse effect on the performance of the decomposition is inevitable. Similar to the Nyquist criteria that requires the sampling rate be at least the twice of the highest frequency included in the signal, for a proper representation of the signal, Rilling stated the minimum requirement for the sampling period to be at most one half the minimum distances between consecutive extrema in the signal to ensure that there is no loss of extrema due to sampling rate [23].

3.4.8 Termination criteria

Huang's initial proposition for the SD was setting a value between 0.2 and 0.3. The rationale for this typical value was again on empirical basis. The two Fourier spectra, computed by shifting only five out of 1024 points from the same data, can have an equivalent *SD* of 0.2-0.3 calculated point-by-point [2]. Therefore, the SD is defined as a metric to measure the squared difference, actually the resemblance, in-between the two successive candidates of a particular IMF. Despite its rigorous view, this metric was later criticized in two folds:

i. Sufficiency of the metric for resemblance, and

ii. Independence of the metric from the definitive conditions for the IMFs. The squared difference might be small, but nothing guarantees that the function will have the same numbers of zero-crossings and extrema [3].

In the mean time, the over-decomposition which was stated as a complication can also be given as an example that motivates derivation of a better termination criterion.

In this context, Huang et al. proposed a second criterion based on the agreement of the number of zero-crossings and extrema. In this criterion, sifting process is proposed to be stopped after *S* consecutive times provided that the numbers of zero-crossings and extrema stay the same and are equal or differ at most by one. The new criterion was respectful to the IMF definition but was bringing a new question that concerned the selection of the *S*-number. As a result of extensive studies, the range of *S*-numbers was advised to be set between 4 and 8 [3].

3.5 Comparison With Fourier and Wavelet Transforms

Historically, Fourier spectral analysis has provided a general method for examining the global energy-frequency distributions and dominated the data analysis efforts since soon after its introduction. As a result, the term "spectrum" has become almost synonymous with the Fourier Transform of the data. Fourier analysis has been applied to all kinds of data despite the crucial restrictions of linearity and stationarity [2].

Then, as its successor, the Wavelet Transform has taken the scene with success stories and has become being referred synonymously with “decomposition”. The youngest of the three, HHT, with still lacking theoretical basis sometimes needs to be incorporated with the former methods to present better solutions. The employment of the other methods can be either for pre-processing or post-processing but synergetic.

Since it has been appreciated that a comparative presentation of the methods may help better understanding, following sections are devoted for this purpose.

3.5.1 Comparison with Fourier Transform

The fundamental differences of the two approaches are as follows:

- The FFT assumes stationarity and linearity of the data wheras the HHT does not require the same assumptions.
- In FFT, the basis for the expansion is global and set a priori whereas in HHT the basis is locally and adaptively imposed by the data.
- The FFT processes amplitude-time data and generates an amplitude-frequency description. But the HHT, processing the same data, generates both time-localized amplitude and frequency information.
- The HHT cannot treat functions such as a single cycle of a square wave as there are no maxima and minima to fit, while the FFT can analyze such a signal [33].
- Since the computational demand of sifting cannot be predetermined, HHT shall be computationally much more demanding than FFT, in general.
- FFT has a much complex structure than HHT because of the requirement to create a Fourier series to fit the signal comprising both the included frequencies as well as the portions where the signal is zero [33]. In conjuntion, a problem of spurious frequencies also arises.

The fundamental differences stated above can easily gain acceptance. However, there are two more features that arguing may pose an opportunity for improved understanding.

- FFT clearly distinguishes the frequencies beyond the cut-off frequency (f_c=*0.67*) where as HHT can not. The HHT generates a single frequency which is equal to the average of the two sine wave frequencies comprising the signal [23, 33].
- As the data length is reduced (smaller number of cycles of the beat pattern are subjected to analysis) so that the signal does not go beyond the first node of the envelope, then the FFT shows only one peak at the average value of the two frequencies [33].

These last two matters will be presented with illustrations in Section 4.5.

3.5.2 Comparison with Wavelet Transform

WT and HHT have both differences and similarities which are as follows:

- Both transforms assumes that a signal is comprised of components that can be decomposed into high and low-frequency components. As the high frequency part is filtered, the decomposition process continues over the low frequency part which is assumed as the new signal.
- In WT, the basis is set a priori and the filters applied along the sifting process are linear time invariant where as in HHT the basis is imposed by the data locally and adaptively.
- The filtering features for WT are given in Figure 3.3 [19]. It can be seen that:
  i. The bandwidth of approximation and detail terms of wavelet decomposition is decided by the sampling frequency of the signal and the level of decomposition instead of the signal characteristics.
  ii. The frequency band of the signal is divided into two equal parts of high frequency and low frequency in every step of wavelet decomposition. Hence, wavelet decomposition is not self-adaptive.
- On the other side, filtering features for HHT is given in Figure 3.4. It can be seen that:
  i. The bandwidth of an IMF is determined by the features of the signal itself rather than other factors, such as sampling frequency as it is in wavelet transform.

ii. There is not a predetermined decimation in the frequency band. The IMF c_1 covers the highest frequency band and the latter IMFs cover lower and lower bands until the residue m indicates the mean trend of the signal. Therefore, EMD is in nature an adaptive band-pass filter bank, whose bandwidth is self-adaptively determined by the analyzed signal. For an extensive analysis of the filtering features of EMD can be referred [19, 23, 29, 34].

- Wavelet decompositions present the advantage of being based on solid and well-understood theoretical foundations, and to be equipped with extremely efficient fast algorithms. This of course contrasts with the present situation of EMD, whose definition is only given as the output of an algorithm, and which clearly lacks from a well-established theory [34].

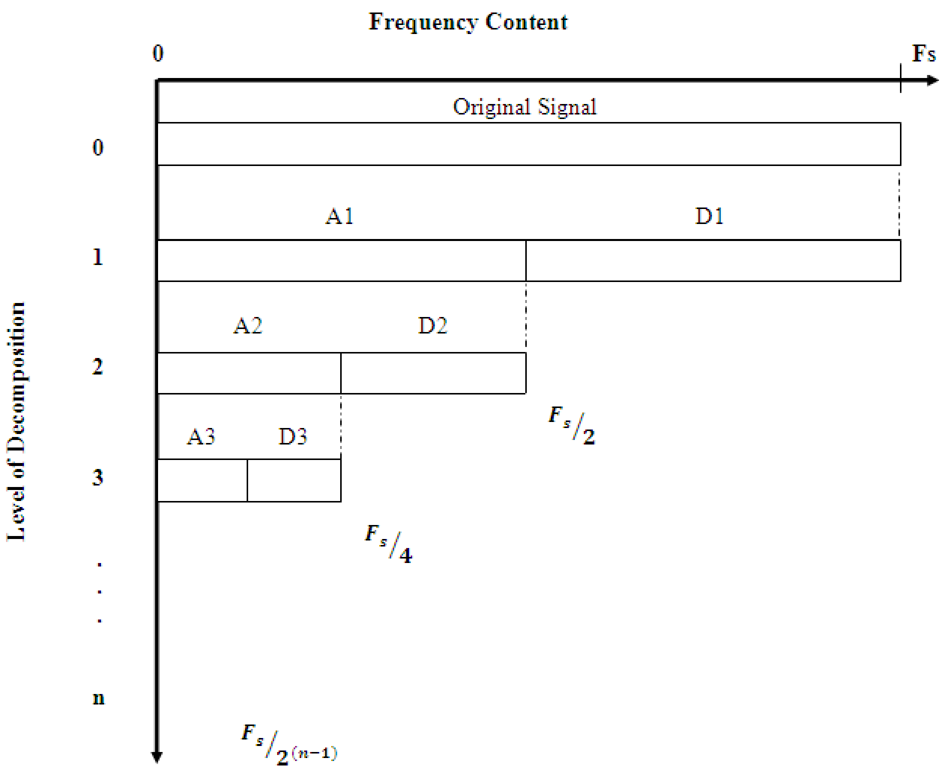

Figure 3.3: The filtering feature of Wavelet Transform.

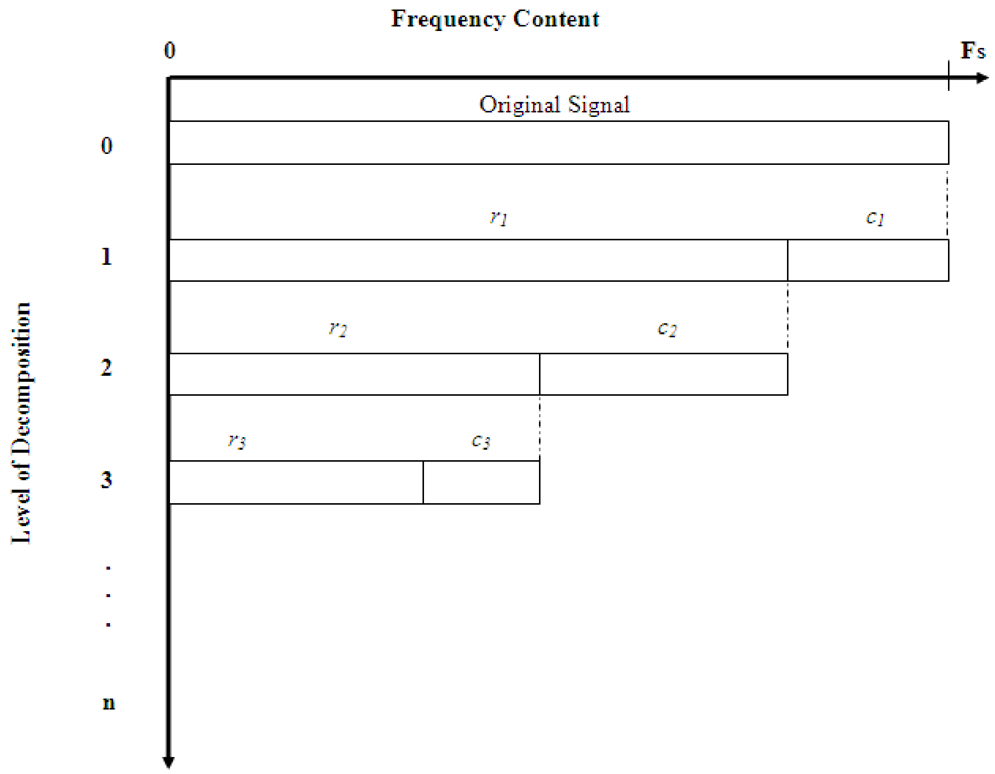

Figure 3.4: The filtering feature of Empirical Mode Decomposition.

A summary of comparison between Fourier, Wavelet and HHT analyses is given in the Table 3.1 [3]:

- **Table 3.1:** Summary of transform properties.

	Fourier	Wavelet	HHT
Basis	a priori	a priori	Adaptive
Frequency	Convolution, Global, Uncertainty	Convolution, Global, Uncertainty	Differential, Local, Certainty
Presentation	Energy-Frequency	Energy-Time-Frequency	Energy-Time-Frequency
Nonlinear	No	No	Yes
Nonstationary	No	Yes	Yes
Feature extraction	No	Yes	Yes
Theoretical base	Complete	Complete	Empirical

3.6 On The Applications to Rotating Machinery and Vibration Signals

The analyses of vibration signals showed that characteristic frequencies that can reveal the state and condition of the systems are included and can be identified. As an example, the bearings, which are the biggest single cause of motor failures [8, 9], generate vibrations of particular frequencies as they get degraded or even they are fault-free [10]. It has been also shown that analytical calculations of the frequencies are also possible by **(3.26)-(3.29)** [8]:

$$f_b = \frac{PD}{BD} f_r \left[1 - \left(\frac{BD}{PD}\right)^2 cos^2(\beta)\right] \tag{3.26}$$

$$f_o = \frac{n}{2} f_r \left[1 - \frac{BD}{PD} cos(\beta)\right] \tag{3.27}$$

$$f_i = \frac{n}{2} f_r \left[1 + \frac{BD}{PD} cos(\beta)\right] \tag{3.28}$$

$$f_t = \frac{1}{2} f_r \left[1 - \frac{BD}{PD} cos(\beta)\right] \tag{3.29}$$

where, n is the number of balls, *PD* is the bearing pitch diameter, *BD* is the ball diameter, β is the contact angle of the balls on the races and f_r is the rotational speed of the rotor in Hz.

The f_o and f_i, for most bearings with number of balls between 6 and 12, can be approximated by **(3.30)** and **(3.31)**:

$$f_o = 0.4 n f_r \tag{3.30}$$

$$f_i = 0.6 n f_r \tag{3.31}$$

This generalized formulation allows an easier definition of frequency bands without requiring explicit knowledge of bearing construction [8].

Despite the analytically developed background, the success of the conventional signal analysis methods in the analysis of vibration data cannot be more than partial because of the general assumption that the process generating the data is stationary and linear. Relevantly, it has been known and shown that the systems were nonlinear due to a number of factors like load, friction/rubbing, uneven clearances, non-linear stiffness terms, etc. and non-stationary due to transients and time varying nature and FFT alone is not capable of analyzing the frequency content of a defective bearing signal [36].

Signals with such non-linearity are best analyzed in time–frequency domain and as the best available time–frequency method so far, Wavelet transform has become a well accepted time–frequency analysis tool and has been widely used in vibration signals analysis. But the wavelet transform still cannot fulfill the rolling bearing fault detection task very well since it has some inevitable deficiencies [31]. The results of wavelet transform depend on the choice of the wavelet basis function. Only signal characters that correlate well with the shape of the wavelet basis function have a chance to lead to coefficients of high value, while all other features will be masked or even completely ignored [19]. Additionally, the interference terms, border distortion and energy leakage, all of which will generate a lot of small undesired spikes all over the frequency scales and make the results confusing and difficult to be interpreted.

Since the frequency ranges of the vibration signals are often wide; implicit constraints of high sampling rate and large data span need to be satisfied requiring that the desired method should have good computational efficiency.

Unfortunately, the computation of Continuous Wavelet Transform (CWT) is somewhat time consuming and is not suitable for large size data analysis. The recent popular time–frequency analysis method, Hilbert Huang Transform (HHT), has good computational efficiency since its most time consuming step, EMD operation, does not involve the convolution and other time-consuming operations, therefore the HHT can deal with the signals of large size [31].

Moreover, the occurrence of impacts in bearings is often rather frequent and the interval between two adjacent impacts is quite small. Hence, the desired time–frequency analysis methods for bearing vibration signal analysis should have fine resolutions both in time domain and in frequency domain. Due to the limitation of Heisenberg–Gabor inequality, the wavelet transform cannot achieve fine resolutions in both time domain and frequency domain simultaneously, therefore, although the wavelet transform has good time resolution in high frequency region, it often cannot separate those impacts, for the time interval between them are often too small. HHT does not involve the concept of the frequency resolution and the time resolution. So the HHT seems to have potential to become a perfect tool for rolling bearing fault detection and vibration data analysis [31].

The implementations of HHT on vibration signals are either with or without modifications on the algorithm. The modifications are also in two folds; with improvements on the HHT algorithm or with incorporation of other methods such as WT, WPT and FFT. For further details concerning the implementations on vibration signals, the references such as [19, 27, 30, 31, 36, 37] which are attributed for a general summary can be referred.

4. THE MODIFIED HILBERT HUANG TRANSFORM

In this chapter, the modified Hilbert Huang Transform is introduced as an improved version of the classical method. The details of the modifications are presented throughout this chapter.

4.1 Objectives

In the previous section, following the introduction of HHT method, features and outstanding problems of the method were mentioned. Of those matters, local mean approximation and the end effect are considered with the objective to propose improvement. As the chapter concludes, it will be shown that the proposals result well with better frequency resolution.

4.2 Local Mean Approximation

In the original method, the local mean approximation is calculated as the local average of the two cubic splines generated with respect to the maxima and minima of the signal. Huang's proposition was to include the first and the last elements of the time series in both the maxima and minima vectors. The visualization of the process was presented in Figure 3.1.

In the alternative method proposed to generate the local mean curve, there exists only one vector that retains the extrema including the first and the last values of the time series. As the extrema are identified, the knots for the spline comprising the local mean approximation are calculated as the average of the two successive extrema and a cubic spline is fit. Prior to the sifting operation:

- The vector retaining the local mean approximation needs to be modified for the end vectors with respect to the criteria stated in Section 4.3, and
- The spline needs to be resampled at the original data locations.

Afterwards, the sifting is realized by computing the difference between the signal and the local mean approximation. The phases of the process is depicted in the Figure 4.1. As it can be verified from the Figure 3.1 and 4.1, the two mean curves obtained with the original and proposed methods resemble each other. The significance of the difference and relation with each other is left as further study subject.

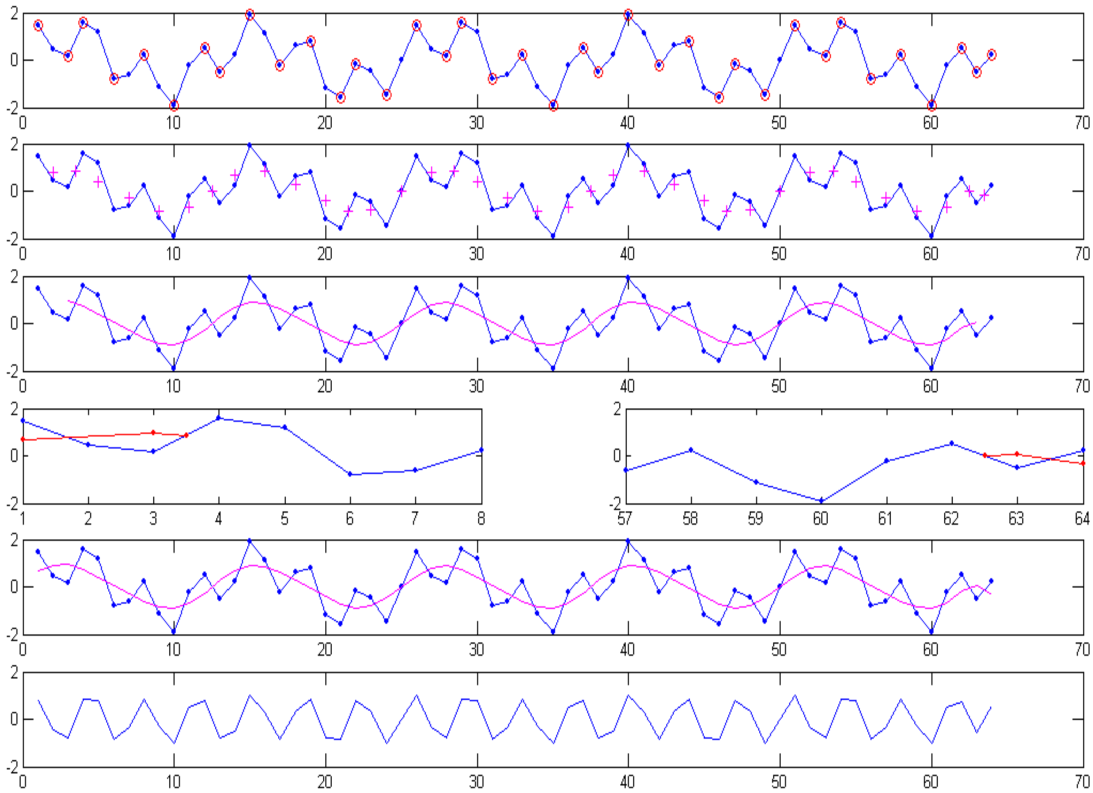

Figure 4.1: Phases of the Modified Algorithm – First iteration.

4.3 Adjustment of the Boundary Values

The first and the last elements of the time series were included in the extrema vector even though they might not be actual extrema. Thus, the part of the spline that resides between the 2^{nd} and $(n-1)^{th}$ extrema location is assumed as is while the tip sections, the prefix and the suffix, are generated with respect to the following formulation and concatenated with the spline.

Adjustment of the Prefix:

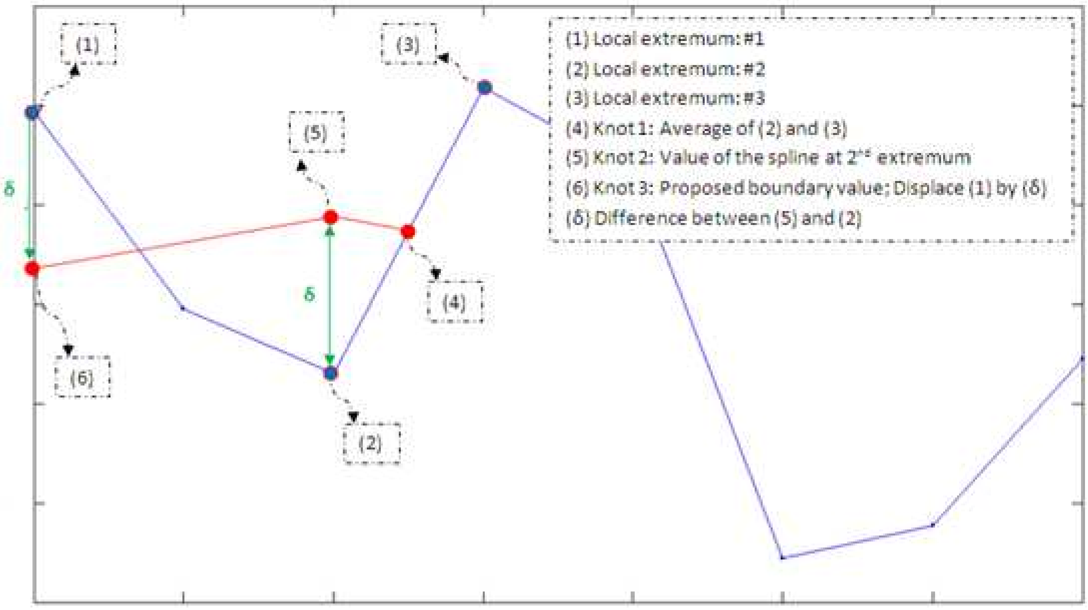

Figure 4.2: Adjustment of the Prefix (fore tip of the time series).

First of the two definitive conditions for IMF requires symmetric wave profile. With this rationale, a point in the δ vicinity of the first element of the time series can be defined. The δ value is the difference between the first actual extrema and corresponding value of the local mean approximation (the spline). So the tip is let to oscillate with an amplitude as the data imposes. The oscillation between the local mean approximation is almost symmetric or eventually gets symmetric. If the number of points before the first actual extrema (node (2) in the figure) is greater than one, then a spline needs to be fit from the points (4), (5) and (6) of Figure 4.2. These points are the knots of the spline and defined as the average of extrema 2-3, value of the spline at 2^{nd} extremum and proposed value obtained by displacement of the first element (first extrema) of the time series by δ, respectively. The same rationale applies for the adjustment of suffix and illustrated in Figure 4.3.

Adjustment of the Suffix:

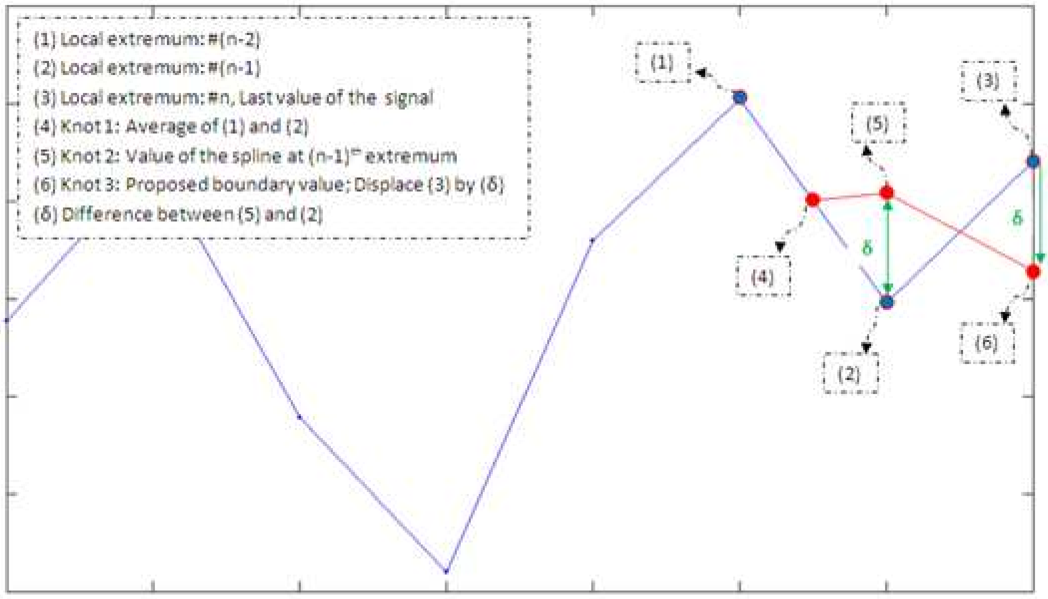

Figure 4.3: Adjustment of the Suffix (rear tip of the time series).

4.4 The Algorithm

(0) Load / Generate the signal "s"

Assign "s" to "y" to be used during the sifting process.

"y" holds the amplitude and "x" holds the location index.

(1) Find extrema (including the first and last values of the signal):

"Ey" holds the amplitude value and "Ex" holds the location index of the extrema .

(2) Compute the coordinates of the knots for the spline: *aveX*, *aveY*

$$aveY(i) = \frac{1}{2}[Ey(i) + Ey(i+1)] \tag{4.1}$$

$$aveX(i) = \frac{1}{2}[Ex(i) + Ex(i+1)] \tag{4.2}$$

(3) Fit cubic spline to the knots

(4) Resample the spline at original data locations excluding the two boundary sections:

Section 1: $[\ Ex(1)\ , Ex(2)\ )$

Section 2: $(\ Ex(n-1)\ , Ex(n)\ ]$ where *n* is the length of the extrema vector

(5) Adjust the boundary sections of the spline

(6) Concatenate the boundary sections with the resampled spline to obtain the approximation vector m

(7) Sifting

$$y = y - m \tag{4.3}$$

Check whether the termination criteria is met or not to decide on the continuation of the sifting. If the criteria is met, assign y as the i^{th} intrinsic mode function (component):

$$c_i = y \tag{4.4}$$

Remove the component c_i from the residual r_i,

$$r_{i+1} = r_i - c_i \tag{4.5}$$

Assign r_{i+1} as the new signal to be sifted,

$$y = r_{i+1} \tag{4.6}$$

(8) The termination criteria can either be selected as the SD value or *S-number* of consecutive sifting iterations. But the application to vibration signals showed that *SD* value should not be the first choice.

4.5 Validation of the Algorithms and Softwares

In accordance with Rilling's study, the signal $x(t)$ is comprised of two superimposed sinusoids.

$$x(t) = cos(2\pi f_1 t) + cos(2\pi f_2 t) \tag{4.7}$$

Equivalently, $x(t)$ may be expressed as [21]:

$$x(t) = 2\,cos\left(\frac{f_1 - f_2}{2}t\right)cos\left(\frac{f_1 + f_2}{2}t\right) \tag{4.8}$$

Mathematically, equation (4.7) resembles a two tone signal whilst equation (4.8) resembles an amplitude modulated single tone signal. Similarly, the physical interpretation of the signal may also be varying. Physically, variation of the interpretation depends on the apartness of the frequencies. As it is illustrated in Figure 4.4, the interpretation shall be a two tone signal if the two frequencies are sufficiently apart and an amplitude modulated single tone signal otherwise. Latter is referred as the "beat effect" implying that the sum of two components yields a third signal without any resemblance of the components [23]. More than being a physical interpretation and a virtual situation, this is what EMD concludes in compliance with Rilling's conclusion, but violating the independent basis paradigm.

In order to illustrate the cases explained above, one being the reference signal three sinusoids are generated as depicted in Figure 4.4. The signal specifications are as follows:

The sampling frequency, f_s, and data length, N, are specified to be 1000 Hz and 600 Hz, respectively.

$$f_r = 55\,Hz\,, \qquad f_1 = 5\,Hz, \qquad f_2 = 45\,Hz\ . \tag{4.9}$$

While specifying the frequencies, following criteria are regarded:

$$f_1 < f_2 < f_r \ll f_s \tag{4.10}$$

$$\frac{f_1}{f_r} < 0.67 \tag{4.11}$$

$$\frac{f_2}{f_r} > 0.67 \tag{4.12}$$

In the Figure 4.4, the 5 Hz component was selected to gain a better visualization opportunity.

Having the beat effect been presented, the performance of the original EMD and modified EMD algorithms will be presented from now on. The applications of the algorithms make use of the above mentioned signal model which is made up of two different sinusoids. With respect to the frequency ratio of the sinusoids and data span, a total number of four cases arise to be analyzed. Particularly, the frequency ratio may be either above or below the cut-off frequency ratio ($f_c = 0.67$) and the number of data samples can be either high or low. In each case, the data is analyzed with both the original and modified EMD (HHT) methods.

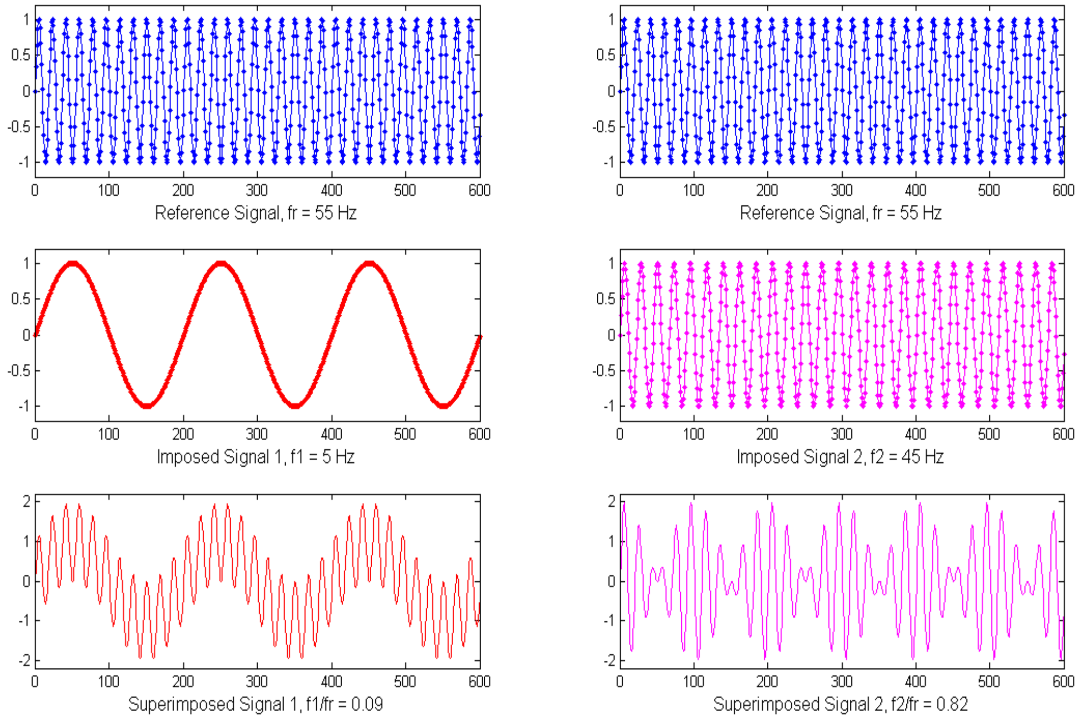

Figure 4.4: The beat effect.

The results concerning each method are presented in two figures:

- *The extracted IMFs and the signal in the time domain.*

  Number of IMFs to be identified is set as four. Each of the identified IMFs are depicted as subplots in red color versus the analyzed signal in blue color.

- *The extracted IMFs and the signal in the frequency domain.*

  The PSD estimates of the analyzed signal and IMFs are given in color coded manner. The matching is as follows: analyzed signal in blue, 1st IMF in red, 2nd in magenta, 3rd in black and 4th in cyan. The justification of using FFT to obtain frequency domain representation of the IMFs will be clarified in Section 5.3.

Brief summary of the application that concerns the four cases corresponding to 8 analyses is given in Table 4.1.

Table 4.1: Frequency resolution analysis summary.

Analysis of the Signal Comprised of Two Sinusoids							
$\frac{f_1}{f_2} < f_c$				$\frac{f_1}{f_2} > f_c$			
64 Samples		1024 Samples		64 Samples		1024 Samples	
Analysis 1 mHHT	Analysis 2 HHT	Analysis 3 mHHT	Analysis 4 HHT	Analysis 5 mHHT	Analysis 6 HHT	Analysis 7 mHHT	Analysis 8 HHT

The results of each decomposition are illustrated in figures that include:

The “extracted IMFs” versus “the signal” in the time domain

The “extracted IMFs” versus “the signal” in the time frequency domain

(Frequency domain representation is obtained by FFT with window size equal to the data span)

Analysis 1: Decomposition with mHHT

Signal Specifications:

$f_1 = 35$ Hz, $f_2 = 55$ Hz. Hence, $\frac{f_1}{f_2} = 0.64 < f_c$. N $= 64$ and $f_s = 1$ kHz.

Analyzed signal (in blue) versus the identified IMFs (in red) are depicted in Fig. 4.5. As it can be verified from Figure 4.6, the PSD estimate of the signal does not help identification of the two components. The PSD estimates of the 4 IMFs reveal 3 components but with shifted central frequencies with respect to the actual components.

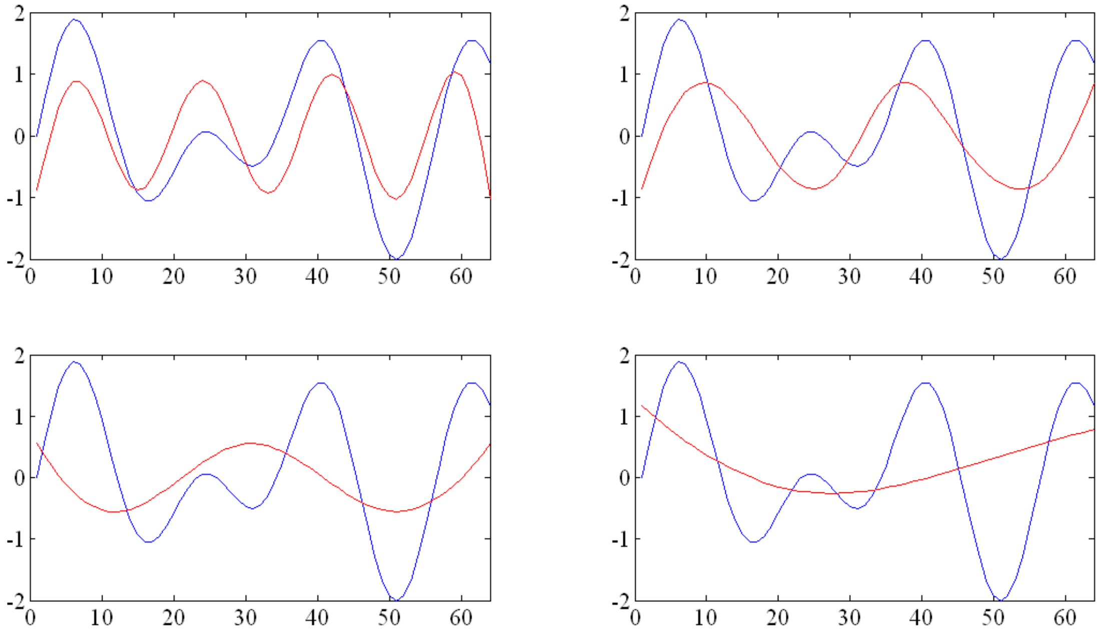

Figure 4.5: The analyzed signal versus identified IMFs in the time domain.

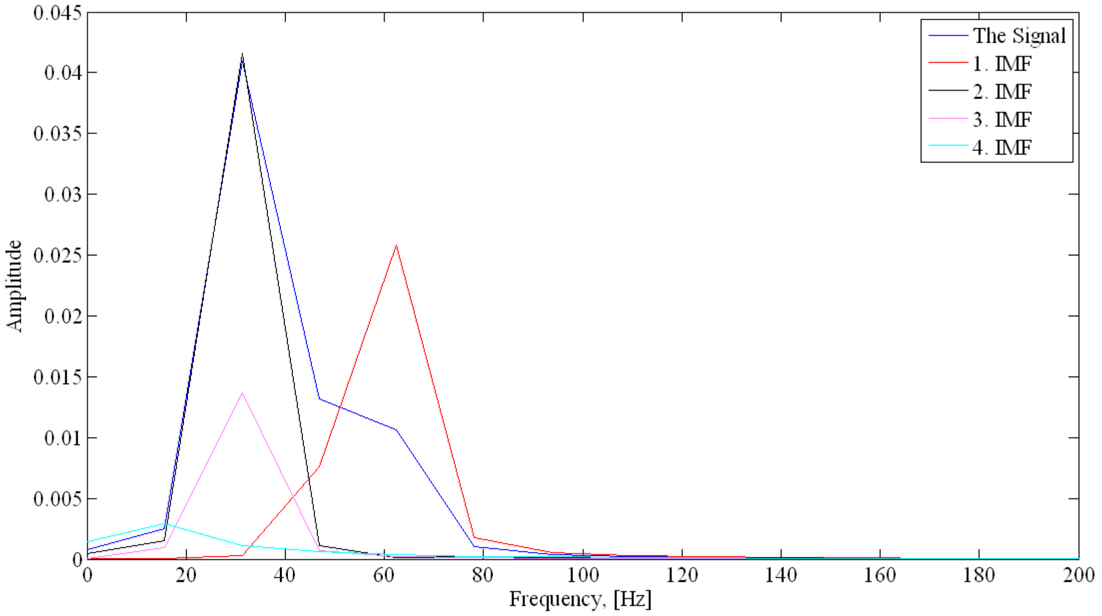

Figure 4.6: The analyzed signal versus identified IMFs in the frequency domain.

Analysis 2: Decomposition with HHT

Signal Specifications:

$f_1 = 35$ Hz, $f_2 = 55$ Hz. Hence, $\frac{f_1}{f_2} = 0.64 < f_c$. $N = 64$ and $f_s = 1$ kHz.

Analyzed signal versus the identified IMFs are depicted in Figure 4.7.

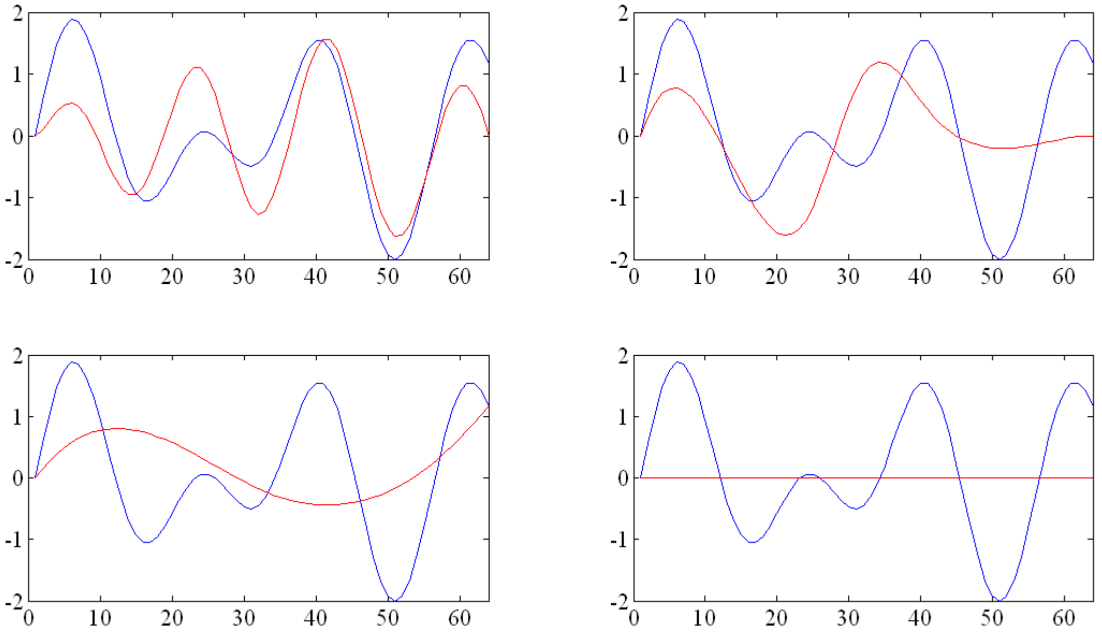

Figure 4.7: The analyzed signal versus identified IMFs in the time domain.

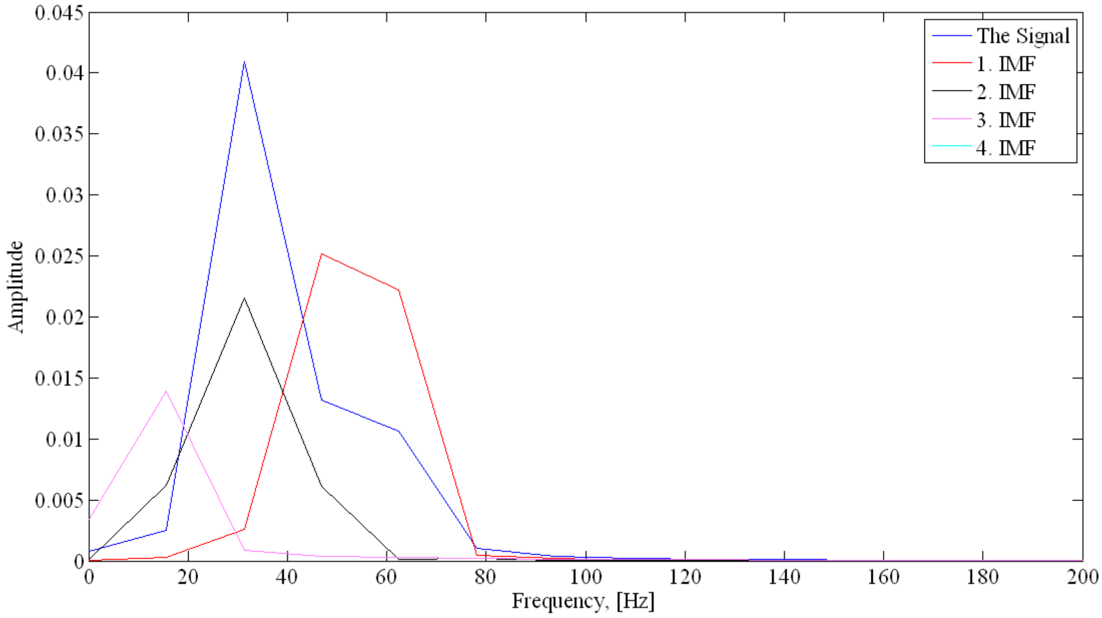

Figure 4.8: The analyzed signal versus identified IMFs in the frequency domain.

As it can be verified from Figure 4.8, the PSD estimate of the signal does not help identification of the two components. The PSD estimates of the 4 IMFs reveal 3 components with shifted central frequencies with respect to the actual components. The bands of the components are wider than the ones in Analysis 1.

Analysis 3: Decomposition with mHHT

Signal Specifications:

$f_1 = 35$ Hz, $f_2 = 55$ Hz. Hence, $\frac{f_1}{f_2} = 0.64 < f_c$. $N = 1024$ and $f_s = 1$ kHz.

Analyzed signal versus the identified IMFs are depicted in Figure 4.9.

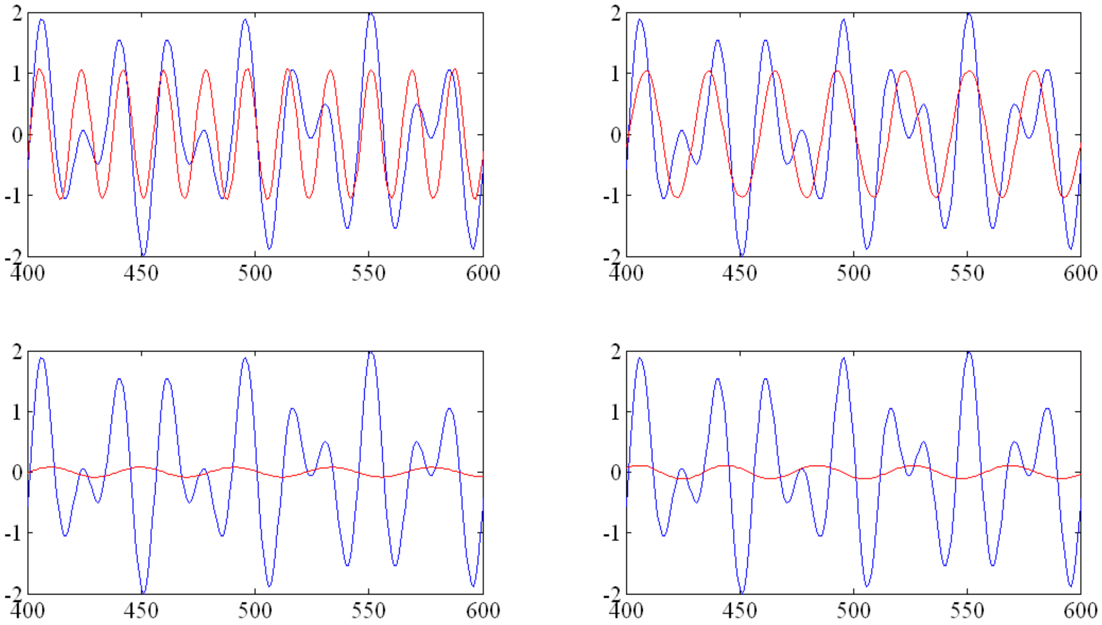

Figure 4.9: The analyzed signal versus identified IMFs in the time domain.

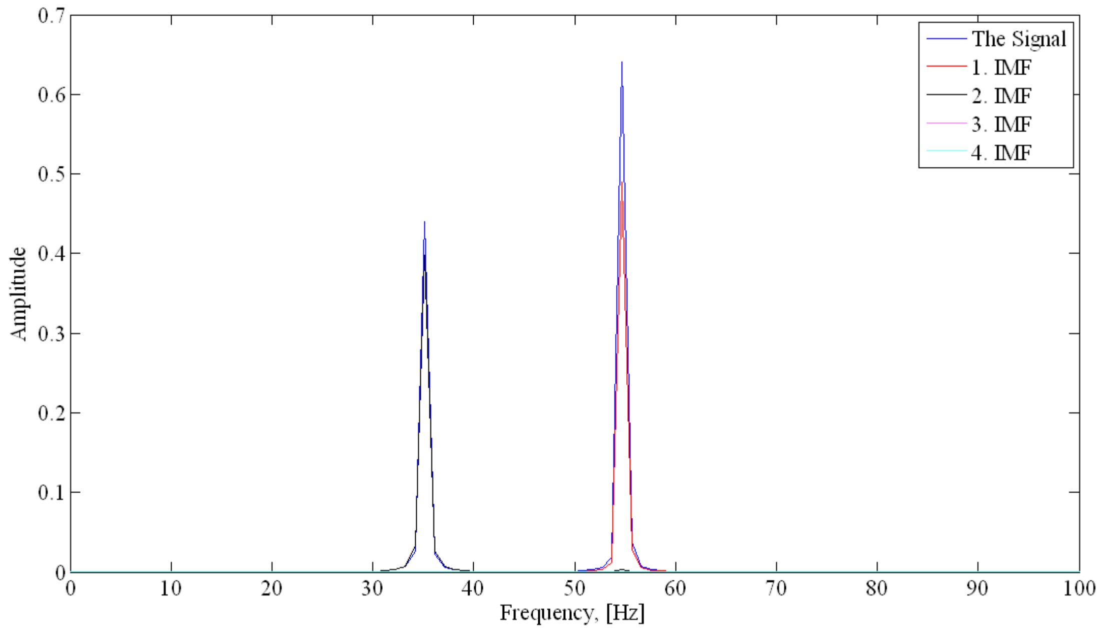

Figure 4.10: The analyzed signal versus identified IMFs in the frequency domain.

As it can be verified from Figure 4.10, the PSD estimate of the signal clearly identifies the two components. PSD estimates of the IMFs reveal that the actual components are resolved separately into the first two of the four IMFs. PSD estimates of them are dominant and coincide with the actual components.

Analysis 4: Decomposition with HHT

Signal Specifications:

$f_1 = 35$ Hz, $f_2 = 55$ Hz. Hence, $\frac{f_1}{f_2} = 0.64 < f_c$. $N = 1024$ and $f_s = 1$ kHz.

Analyzed signal versus the identified IMFs are depicted in Figure 4.11.

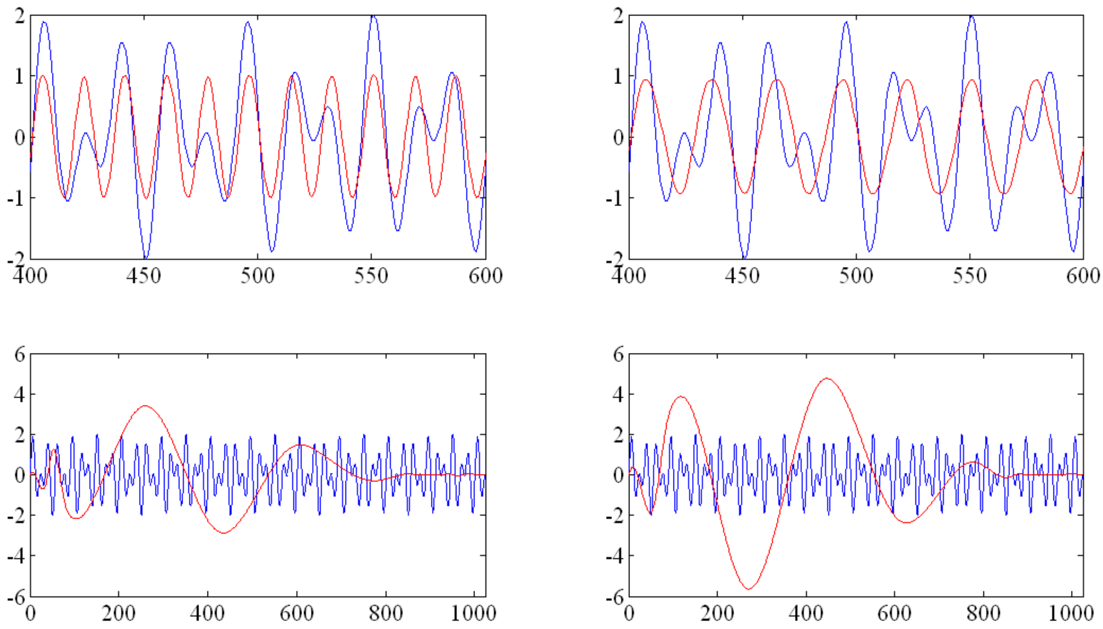

Figure 4.11: The analyzed signal versus identified IMFs in the time domain.

As it can be verified from Figure 4.12, the PSD estimate of the signal identifies four components. PSD estimates of the IMFs reveal that the actual components are resolved separately into the first two of the four IMFs. But the PSD estimates of meaningless low frequency components are dominant..

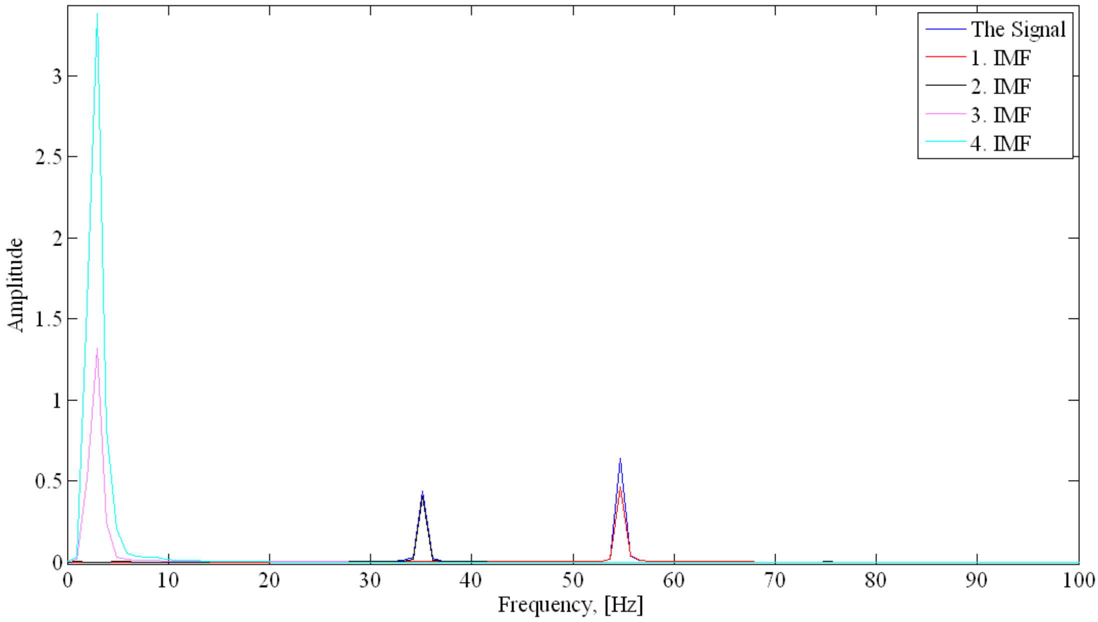

Figure 4.12: The analyzed signal versus identified IMFs in the frequency domain.

Analysis 5: Decomposition with mHHT

Signal Specifications:

$f_1 = 45$ Hz, $f_2 = 55$ Hz. Hence, $\frac{f_1}{f_2} = 0.82 > f_c$. $N = 64$ and $f_s = 1$ kHz.

Analyzed signal versus the identified IMFs are depicted in Figure 4.13.

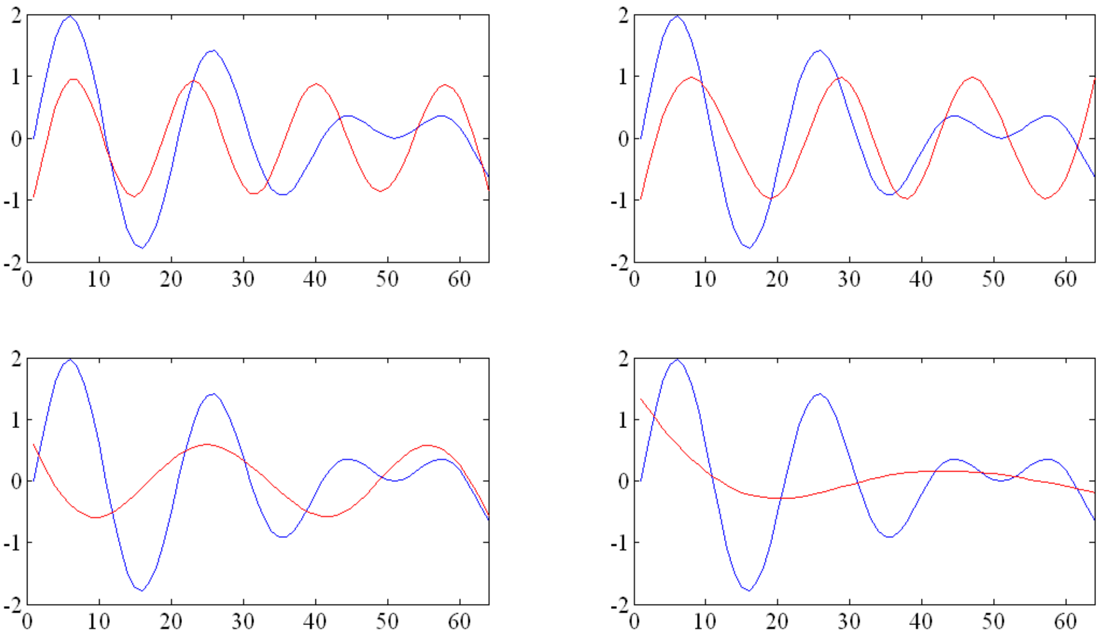

Figure 4.13: The analyzed signal versus identified IMFs in the time domain.

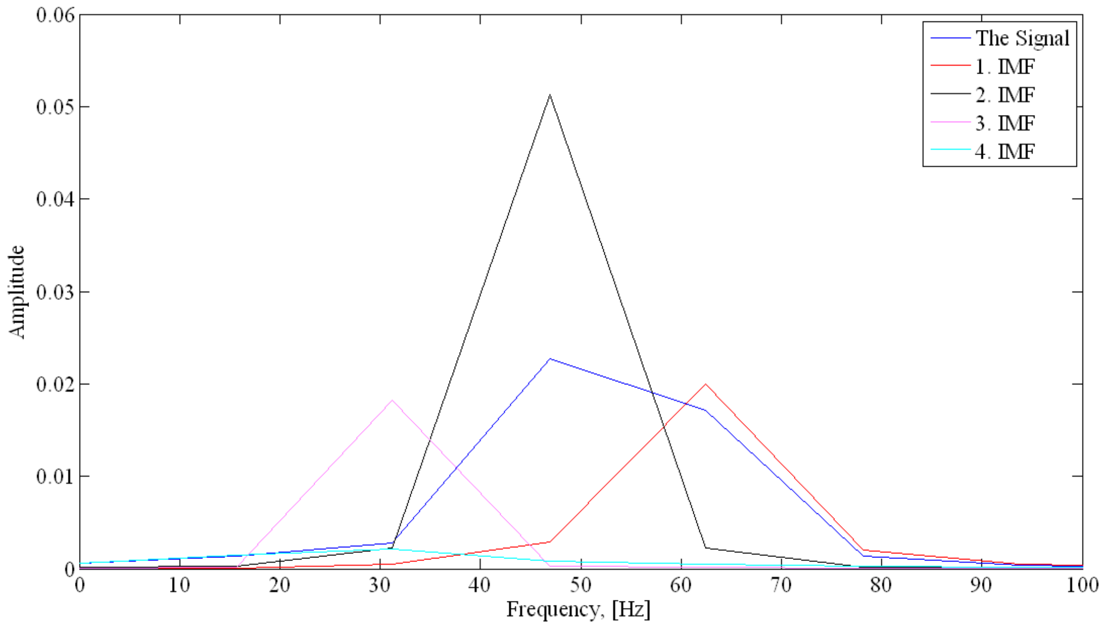

Figure 4.14: The analyzed signal versus identified IMFs in the frequency domain.

As it can be verified from Figure 4.14, the PSD estimate of the signal does not help identification of the two components. The PSD estimates of the 4 IMFs reveal 3 components but with shifted central frequencies with respect to the actual components. The dominant component's central frequency coincides with the average value of the actual components.

Analysis 6: Decomposition with HHT

Signal Specifications:

$f_1 = 45$ Hz, $f_2 = 55$ Hz. Hence, $\frac{f_1}{f_2} = 0.82 > f_c$. $N = 64$ and $f_s = 1$ kHz.

Analyzed signal versus the identified IMFs are depicted in Figure 4.15.

As it can be verified from Figure 4.16, the PSD estimate of the signal does not help identification of the two components. The PSD estimates of the 4 IMFs reveal 4 components with shifted central frequencies with respect to the actual components. The dominant component's central frequency coincides with the average value of the actual components.

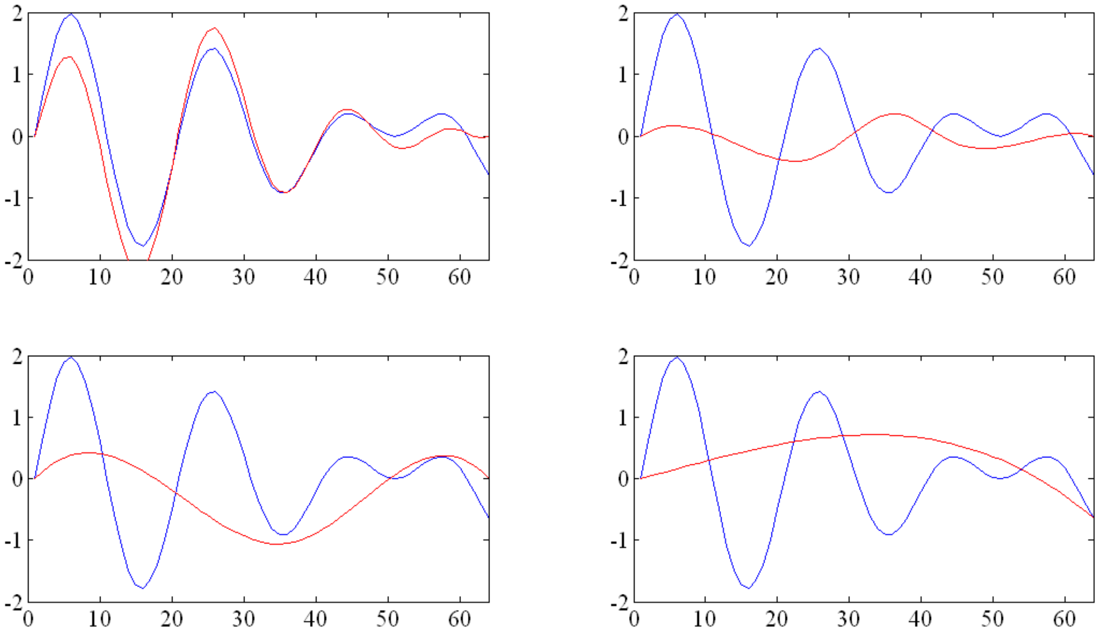

Figure 4.15: The analyzed signal versus identified IMFs in the time domain.

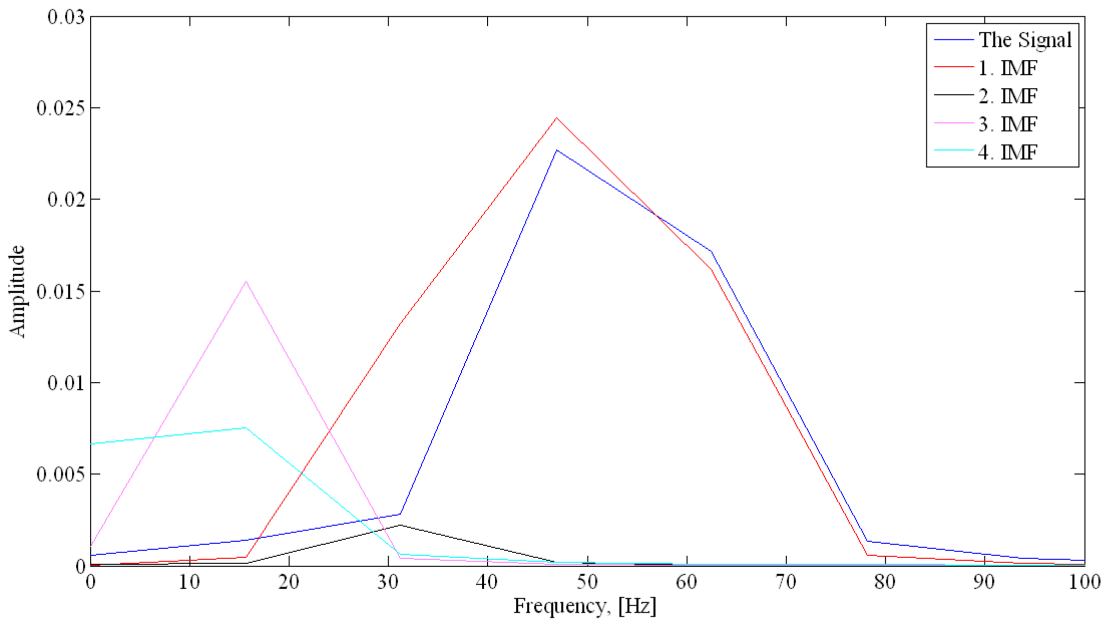

Figure 4.16: The analyzed signal versus identified IMFs in the frequency domain.

Since none of the methods produced meaningful results in Analysis 5 and 6, they are repeated with same specifications except the sample size which is taken as $N=128$.

Repetitive Analysis 5: Decomposition with mHHT

Analyzed signal versus the identified IMFs are depicted in Figure 4.17.

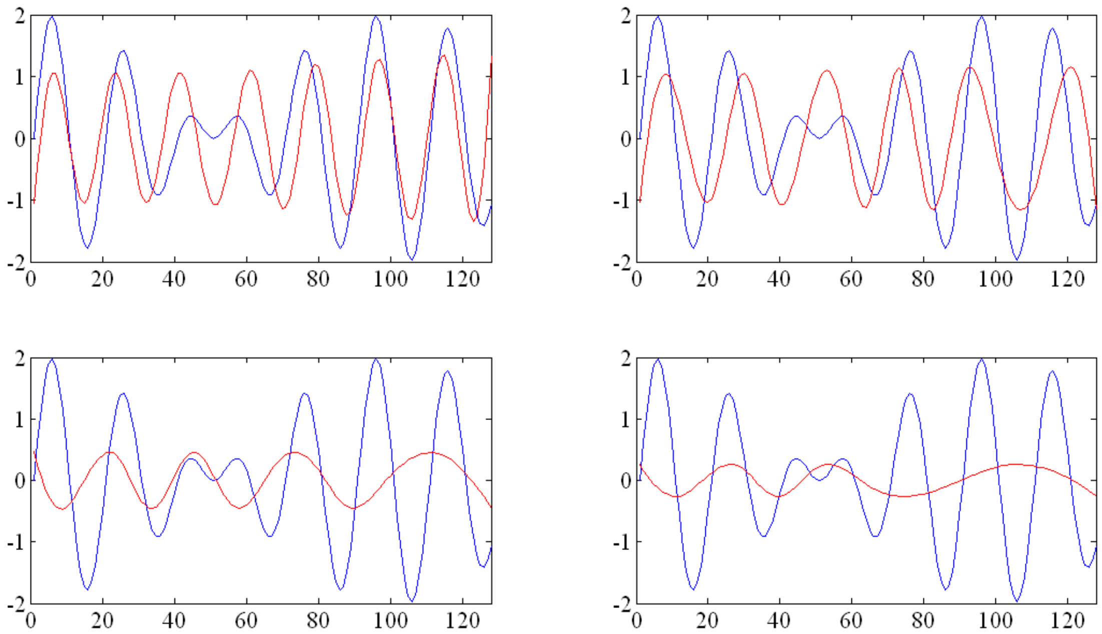

Figure 4.17: The analyzed signal versus identified IMFs in the time domain.

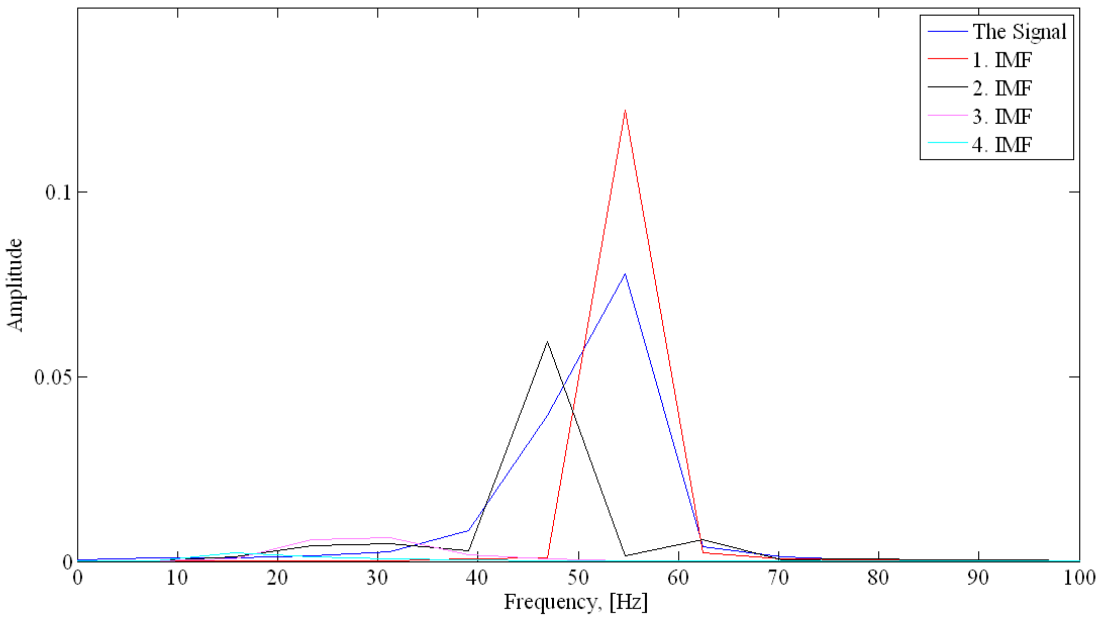

Figure 4.18: The analyzed signal versus identified IMFs in the frequency domain.

As it can be verified from Figure 4.18, the PSD estimate of the signal cannot identify the two components. PSD estimates of the IMFs reveal that the actual components are resolved separately into the first two of the four IMFs. PSD estimates of them are dominant with coinciding central frequencies with the actual components.

Repetitive Analysis 6: Decomposition with HHT

Analyzed signal versus the identified IMFs are depicted in Figure 4.19.

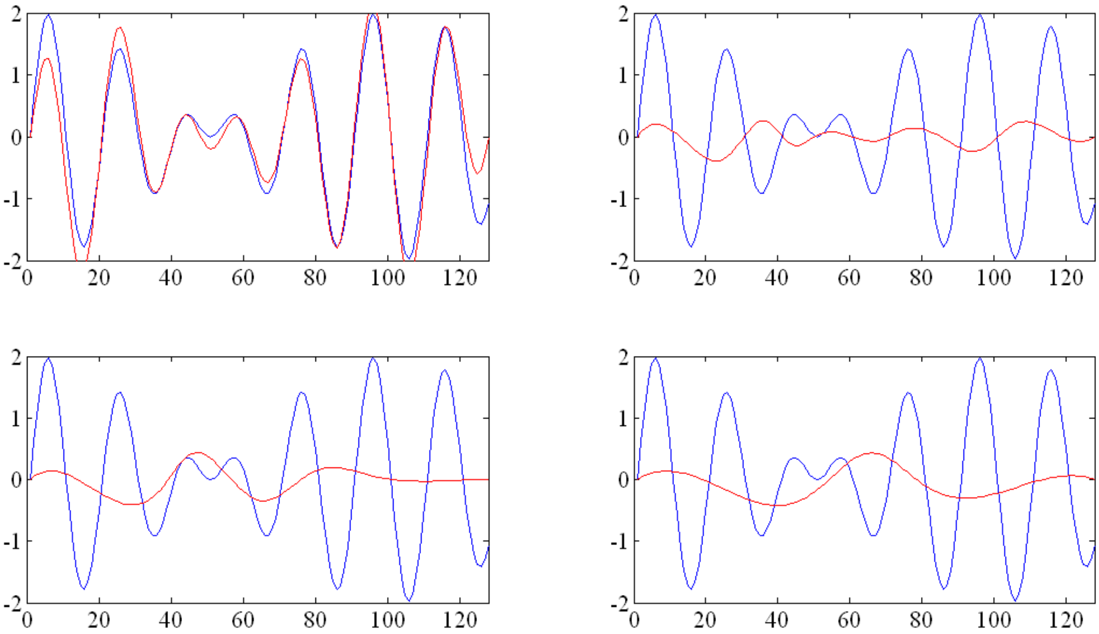

Figure 4.19: The analyzed signal versus identified IMFs in the time domain.

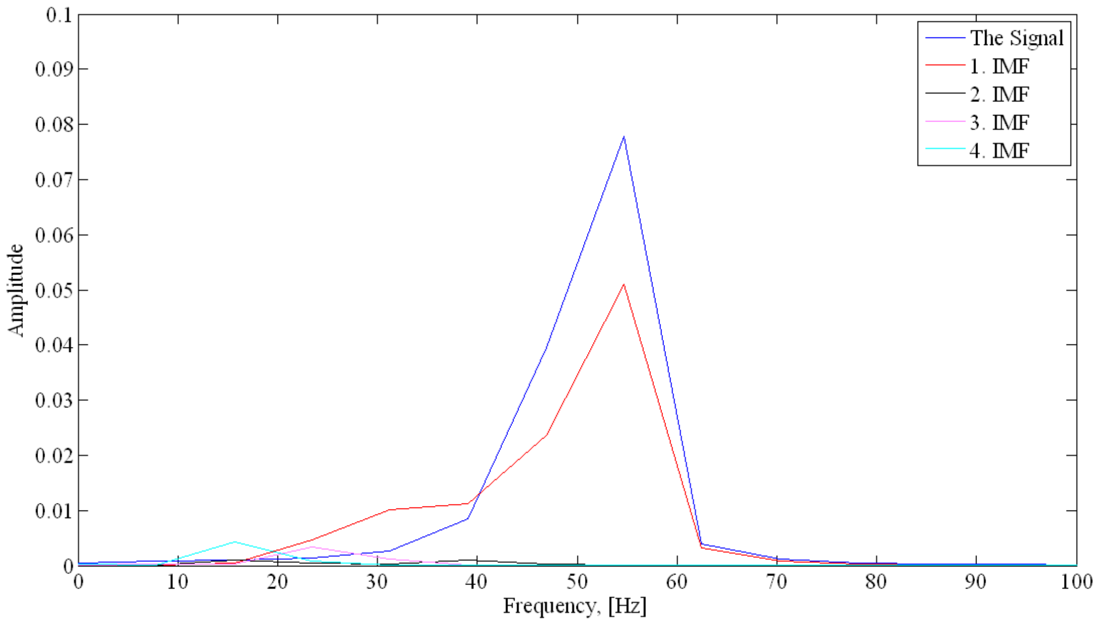

Figure 4.20: The analyzed signal versus identified IMFs in the frequency domain.

As it can be verified from Figure 4.20, the PSD estimate of the signal cannot identify the two components. PSD estimates of the IMFs do not provide a better information than the original signal's PSD estimate.

Analysis 7: Decomposition with mHHT

Signal Specifications:

$f_1 = 45$ Hz, $f_2 = 55$ Hz. Hence, $\frac{f_1}{f_2} = 0.82 > f_c$. $N = 1024$ and $f_s = 1$ kHz.

Analyzed signal versus the identified IMFs are depicted in Figure 4.21.

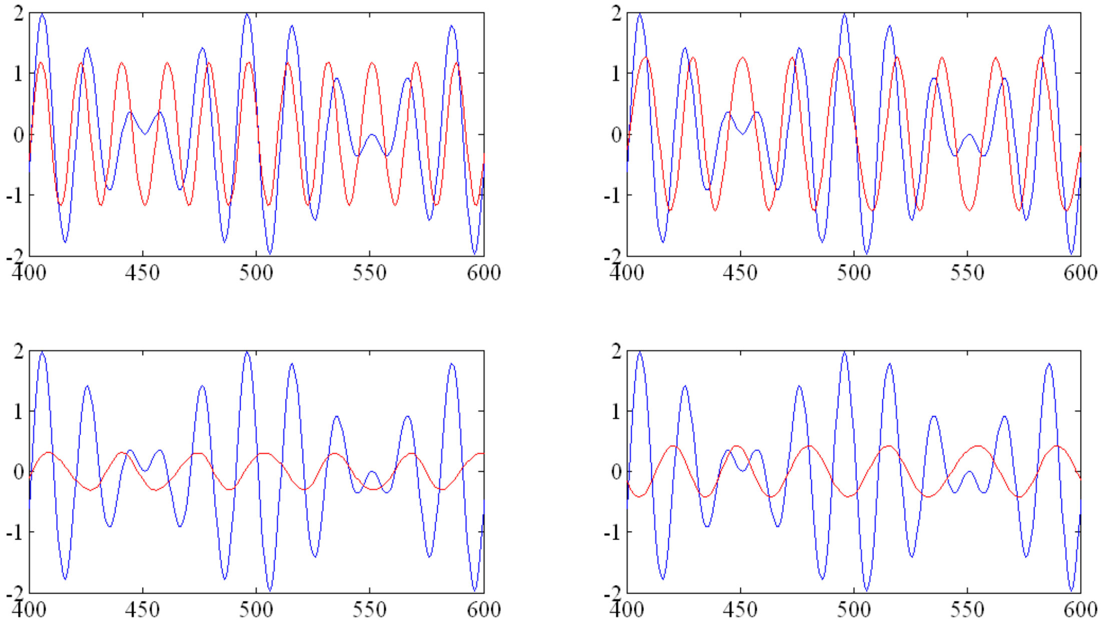

Figure 4.21: The analyzed signal versus identified IMFs in the time domain.

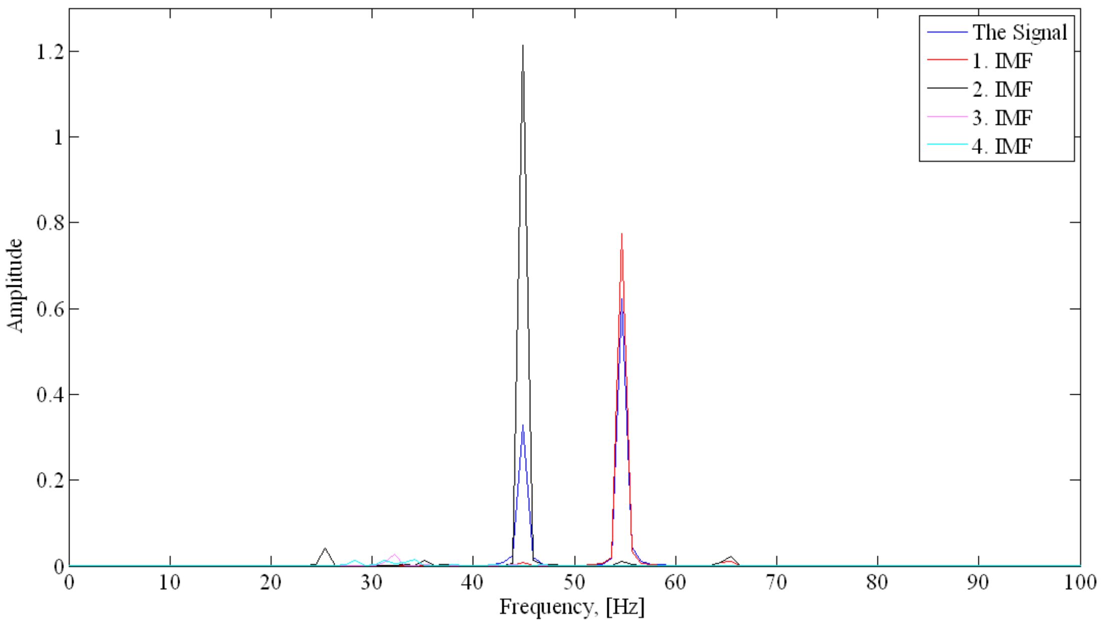

Figure 4.22: The analyzed signal versus identified IMFs in the frequency domain.

As it can be verified from Figure 4.22, the PSD estimate of the signal clearly identifies the two components. PSD estimates of the IMFs reveal that the actual components are resolved separately into the first two IMFs. PSD estimates of them are dominant with coinciding central frequencies with the actual components.

Analysis 8: Decomposition with HHT

Signal Specifications:

$f_1 = 45$ Hz, $f_2 = 55$ Hz. Hence, $\frac{f_1}{f_2} = 0.82 > f_c$. $N = 1024$ and $f_s = 1$ kHz.

Analyzed signal versus the identified IMFs are depicted in Figure 4.23.

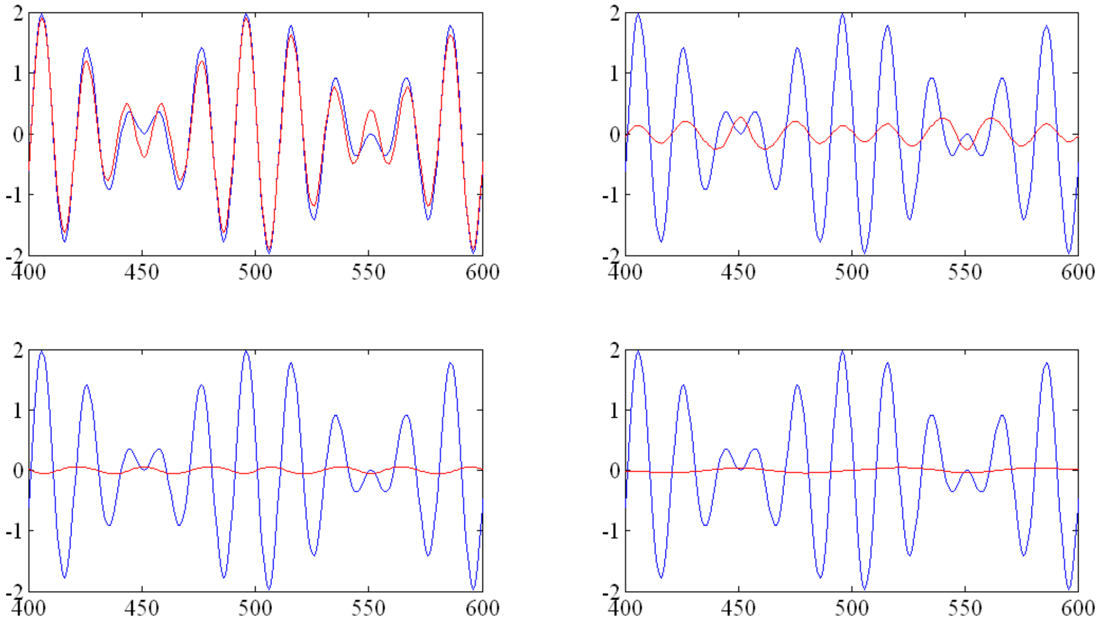

Figure 4.23: The analyzed signal versus identified IMFs in the time domain.

As it can be verified from Figure 4.24, the PSD estimate of the signal clearly identifies the two components. PSD estimates of the IMFs reveal that the actual components are not resolved separately and the first IMF includes both frequencies as a multi-component signal.

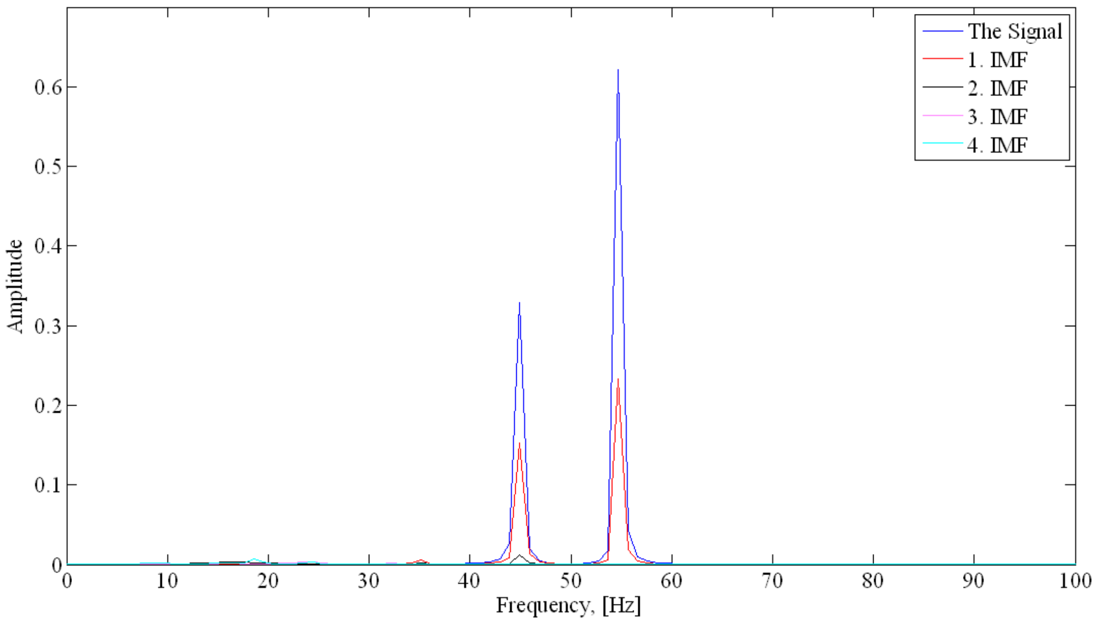

Figure 4.24: The analyzed signal versus identified IMFs in the frequency domain.

4.6 Features / Impacts of the Proposed Method

The features of the proposed method include:

- A new formulation for calculation of the pivot points of local mean representing spline,
- A new formulation for resolution of the boundary effects problem.

The major impacts of the proposed process are:

- Increased resolution in identification of intrinsic frequencies,
- Reduction in number of iterations to identify the intrinsic frequencies.

5. ANALYSIS OF VIBRATION SIGNALS: THE APPLICATIONS

The data consists of the vibration signals acquired from an induction motor during an accelerated aging test that had been performed according to IEEE-Std 117 (1974) test procedures. The details of the test setup can be referred from [5, 6].

5.1 The Data and Its Statistical Properties

The extracts from the time series belonging to the initial and last cycles of the aging process are depicted in Figure 5.1 and related histograms in Figure 5.2.

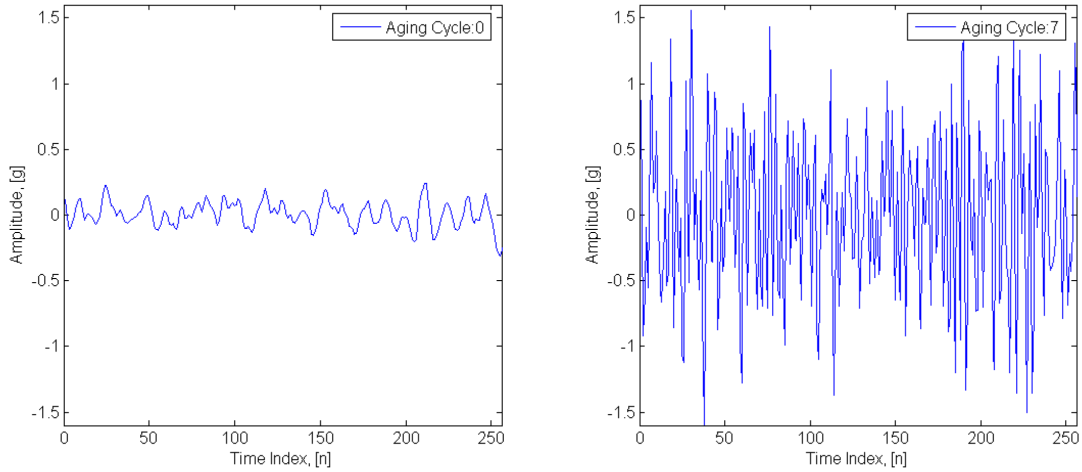

Figure 5.1: Sample sections of data belonging to the initial and aged cycles.

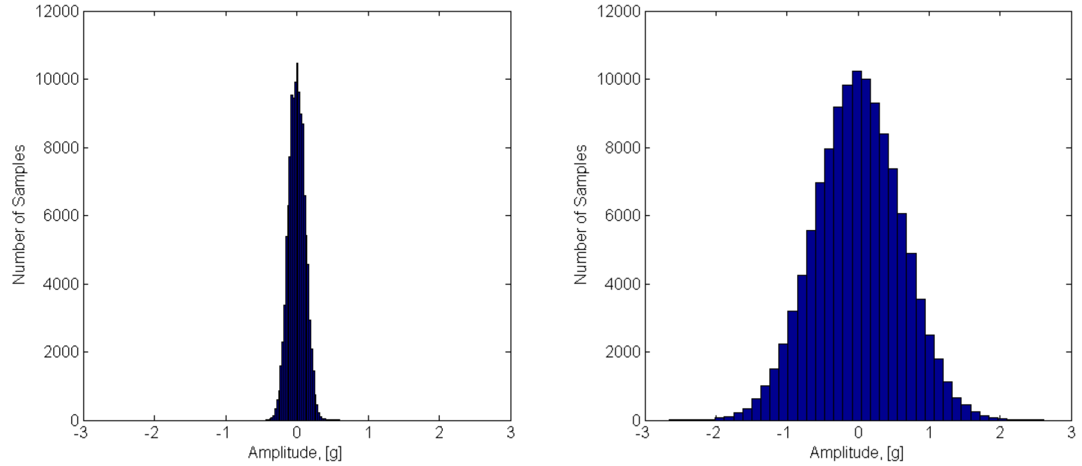

Figure 5.2: Histograms of data belonging to the initial and aged cycles.

The statistical properties of the time series are tabulated in Table 5.1.

- **Table 5.1:** Statistical descriptors of the vibration data.

Cycle of Aging	Mean	Standard Deviation	Skewness	Kurtosis
0	0.0016	0.1135	0.0591	2.9266
1	0.0014	0.1553	-0.0107	3.2875
2	0.0012	0.2082	0.0178	3.0367
3	0.0013	0.2894	0.0162	2.9985
4	0.0009	0.3275	-0.0031	3.0409
5	0.0010	0.3548	-0.0019	2.9832
6	0.0005	0.4147	-0.0103	2.9871
7	0.0030	0.6040	-0.0060	3.0093

In this application the motor speed is 1742 RPM ($\mathrm{f_r} = 29.03$ Hz) and the sampling frequency is 12000 Hz ($\mathrm{F_s} = 12$ kHz). The bearing is of SKF 6205 type.

The approximate characteristic frequencies calculated using equations (3.26)-(3.31) are as follows:

The ball pass frequency of the outer race, $f_o = 104.5$ Hz,

The ball pass frequency of the inner race, $f_i = 156.7$ Hz,

The ball spin frequency, $f_b = 136.9$ Hz.

The fundamental train frequency, $f_t = 11.5$ Hz.

5.2 Analysis of the Data with Conventional Methods

In this section, previous studies with conventional methods such as Fourier Transform and Wavelet Transform are presented.

5.2.1 Fourier spectrum of the vibration data

To extract the variation in the frequency content of the signals in parallel to the aging process cycles, the Fourier transformation is performed and the power spectral densities of these signals are plotted as in Figure 5.3. Comparative analysis of these figures reveals the rise of high frequency components in the vibration signal, between 2-4 kHz, due to accelerated aging process that targets the bearings [8].

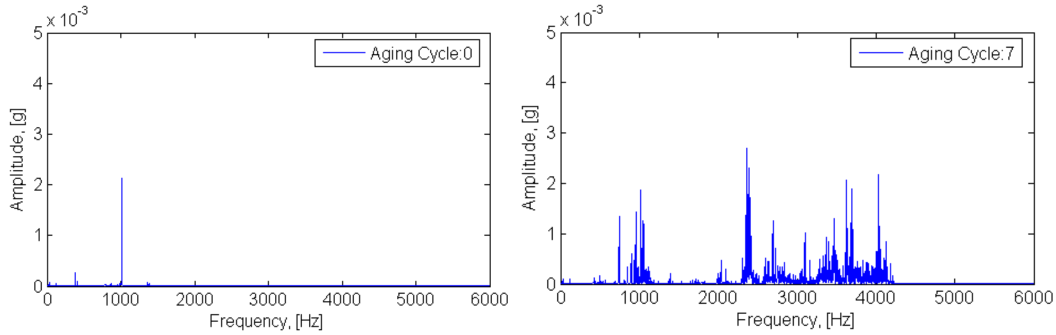

Figure 5.3: PSD of the vibration signals at initial and aged cycles.

The high peaks which are identifiable in the PSD estimates have been known and shown to be the harmonics of the characteristic frequencies like $k_1 f_o$, $k_2 f_i$, $k_3 f_b$ where k_i are integers [8]. At this point, the matter of concern is a priori detection of the potential frequency components that may point and lead to a damage. The continuous wavelet transform, subject of the next section, will be employed to reveal the potential frequency components using only healthy case data.

5.2.2 Wavelet analysis of the vibration data

The continuous wavelet transform produces the different scales and with this way, the small effects of the bearing damage, which can be represented as a prior knowledge between 2-4 kHz defined in high frequency region of the spectrum calculations, can easily be detected. Here, the frequency band of bearing damage which defined as the prior knowledge is a known feature through the related literature. In general, it is accepted as beyond of the 2 kHz. In this sense, the following paragraph gives the details of the continuous wavelet transform application to vibration signals both initial (healthy) and aged (faulty) cases. Continuous wavelet transform is applied to the vibration signals to see how scale or frequency content of a signal changes in time. These signals were decomposed into 256 scales using mother wavelet type of symlet.

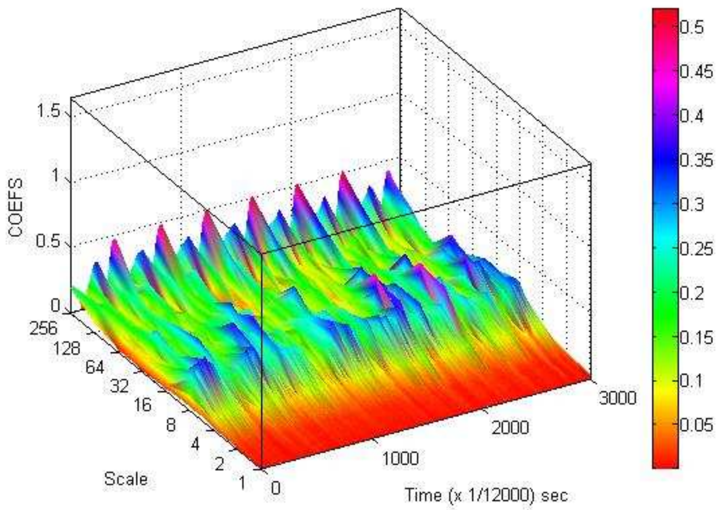

Figure 5.4: Absolute values of CWT coefficients for initial cycle data.

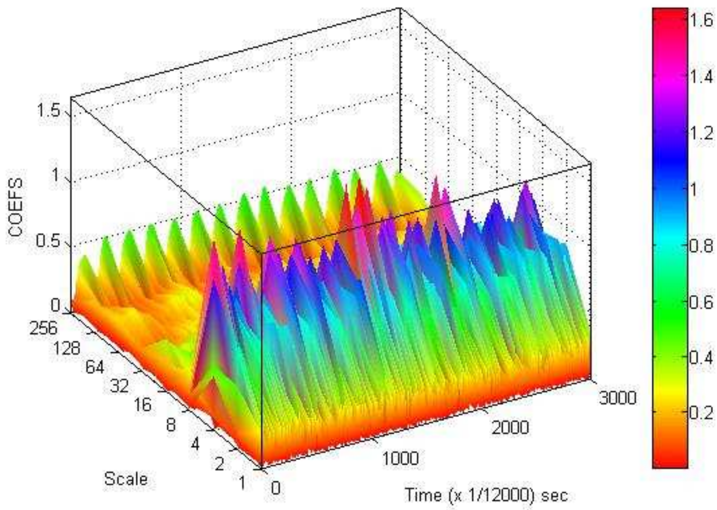

Figure 5.5: Absolute values of CWT coefficients for aged cycle data.

The CWT coefficients are shown in 3D-plane for healthy and aged bearing conditions by means of the Figure 5.4 and 5.5. Comparing these figures it can be observed that bearing damage effect shows itself at low scales as seen in Figure 5.5. Hence, while it can be observed that the rare variations are at the low scale values, high scale values as shown in Figure 5.4 and 5.5 indicate some periodical variations at low frequencies.

However, Time-Scale plot of the Figure 5.4 which is given in Figure 5.6, shows more clearly the small amplitudes of the high frequency components which occur at low scales for determining the potential existence of the bearing damage from the healthy case.

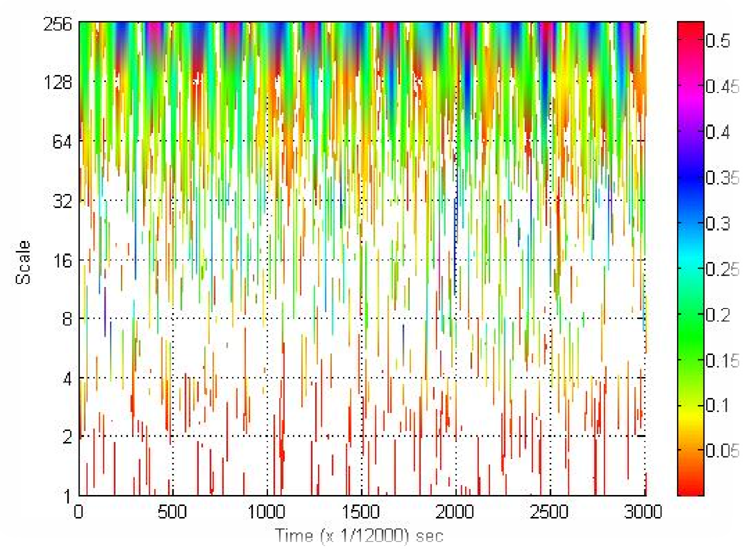

Figure 5.6: Time-Scale plot of the 3D-CWT of the vibration signal in healthy case.

Thus, it can be easily seen the rare small amplitudes of the bearing damage at low scale values, that's high frequency components. These rare small amplitudes represents the primitive effects of the bearing damage effects occurred at high frequency region of the spectrum of the faulty or aged case data. Consequently, the existence of the small amplitude effects related to bearing damage verifies our thesis about this matter. This finding is based on the power of the continuous wavelet transform with its zoom capability and was presented in [8].

In summary, as an alternative method, Continuous Wavelet Transform (CWT) technique is purposed as a new possibility in separating the signals to different scales in terms of using its high capability in revealing the small effects which cannot be noticed on the overall spectrum. Hence, type of very small amplitudes located at high frequency region of the overall spectrum of the healthy case data can easily be extracted to indicate features at low scales of the CWT. This is one advantage of the CWT approach used in this research over the classical Spectral Analysis Techniques. Also these methods create a new possibility to reveal the defective case which causes to the damaged case. Hence it is easily determined by the scaling approach of the CWT technique that magnifies these defective features. This is extremely powerful approach to get the defects in early case and small effect regarding to the bearing damage that denoted between 2-4 kHz are observed by the CWT approach. Therefore, this study indicates an advantage of CWT over the Spectral Analysis Techniques in terms of getting more sensitive fault detection and hence, this possibility can be represented as an important contribution for this study.

Nevertheless, CWT is assessed as a computationally time consuming and unsuitable method for large size data analysis [31]. In this circumstance, search for an computationally efficient alternative method becomes the matter of concern and HHT, which is dealt in the next section, is considered as the alternative.

5.3 Analysis of the Vibration Data with Hilbert Huang Transform

In Chapter 3, it was explained that the HHT, consisting of Empirical Mode Decomposition (EMD) and Hilbert Transform (HT) phases, did not require a priori basis selection and was computationally efficient. However, due to inefficiencies encountered during practical applications, variants of the method are proposed. The effectiveness of utilizing FFT on IMFs from EMD process is reported in [36]. In addition, in [38], the performances of CWT and HHT are evaluated and CWT is reported to produce lower variance frequency estimates than HHT consisting of EMD and HT. In this perspective, corporation of the original EMD algorithm with FFT is preferred for the analysis presented in this section.

The IMFs and PSD of the whole decomposition processes for the initial (0^{th}) and last (7^{th}) cycles are provided in the Appendix A.1-4. In this section, only the data concerning these cycles of the aging process will be presented with the limitation for the IMFs to retain information particular to 2-4 kHz region. Hence, the original results of CWT may serve as a reference for benchmarking and results can be presented without complication of plenty of figures.

The IMFs and corresponding PSD for the 0^{th} and 7^{th} cycles are depicted in Figure 5.7-5.10. In general, from the figures, it can be verified that each IMF retain particular information corresponding different frequency regions in the spectrum and the first IMF includes the distinguishable highest frequencies. Particularly, this remark cannot be precisely verified from Figure 5.7 as long as the frequency interval of [1500 Hz, 4500 Hz] is not zoomed in as it is depicted in Figure 5.8.

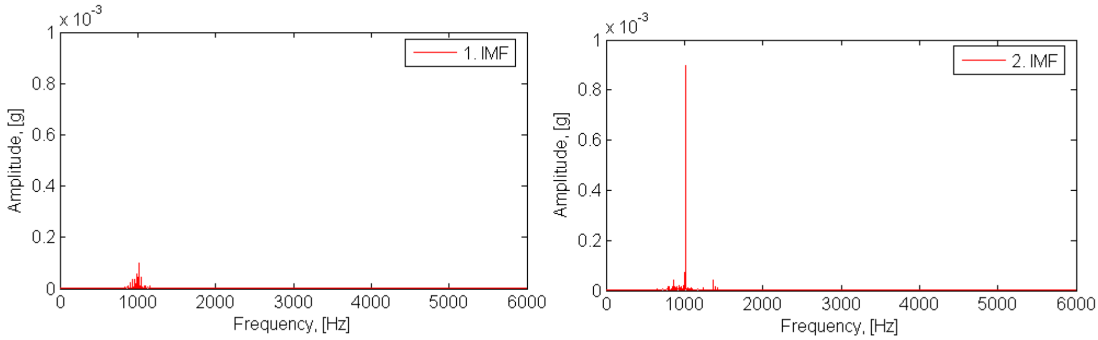

Figure 5.7: PSDs of the first two IMFs of the initial cycle data.

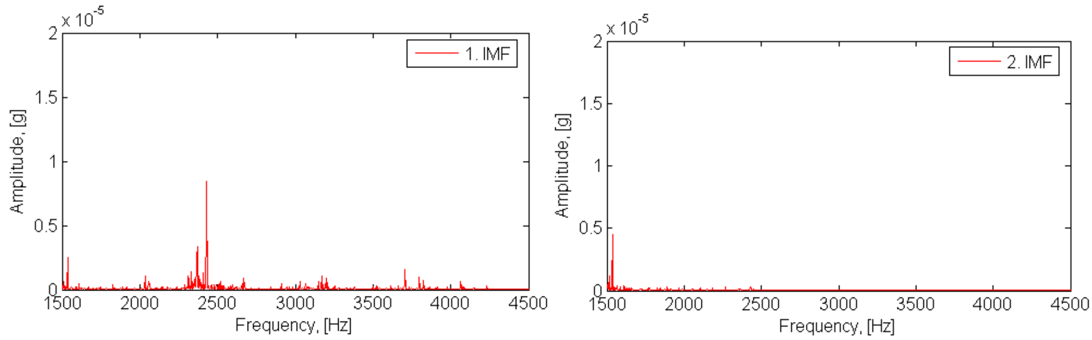

Figure 5.8: Zoomed in values for Figure 5.7.

From the comparative analysis of PSD estimates in Figure 5.8, it can be concluded that the first IMF retains characteristic frequencies between 1.5-4 kHz but initially with small amplitudes. The second IMF's content corresponds to 0.8-1.5 kHz band and the high frequency components of the first IMF are filtered.

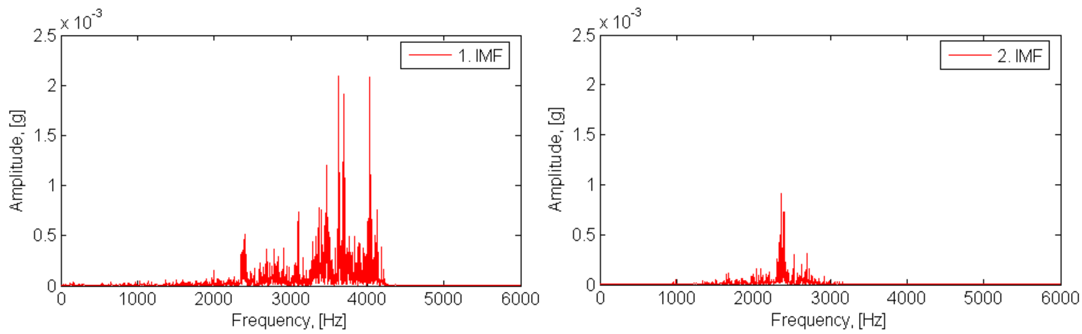

Figure 5.9: PSDs of the first two IMFs of the last cycle data (aged case).

In the initial cycle, the 0.8-1.5 kHz band was retained by the second IMF. However, as it can be verified from Figure 5.9 and 5.10, in the aged case the same band is retained by the 3rd and 4th IMFs. This situation implies that new components with higher frequencies are identified in accordance with expectations.

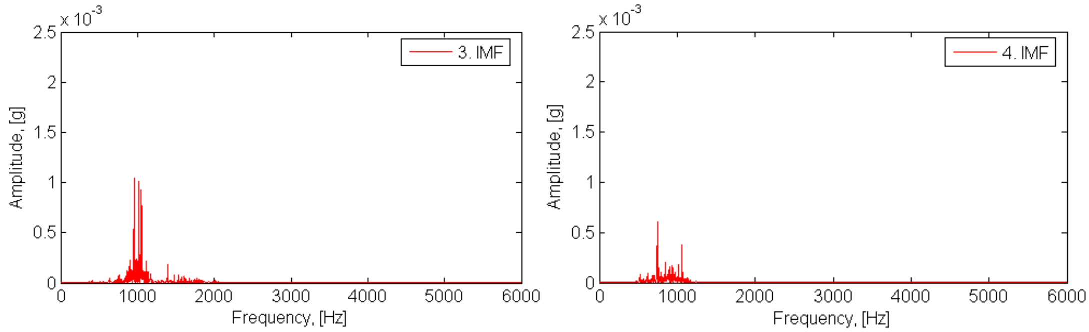

Figure 5.10: PSDs of the 3rd and 4th IMFs of the last cycle data (aged case) extracted by HHT.

Comparison of the PSD estimates of the IMFs of each aging cycle reveals the rise of new components and the strengthening of existing ones in the high frequency region. A closer view of the high frequency content of the first IMFs' for the initial and aged cycles are depicted in Figure 5.11.

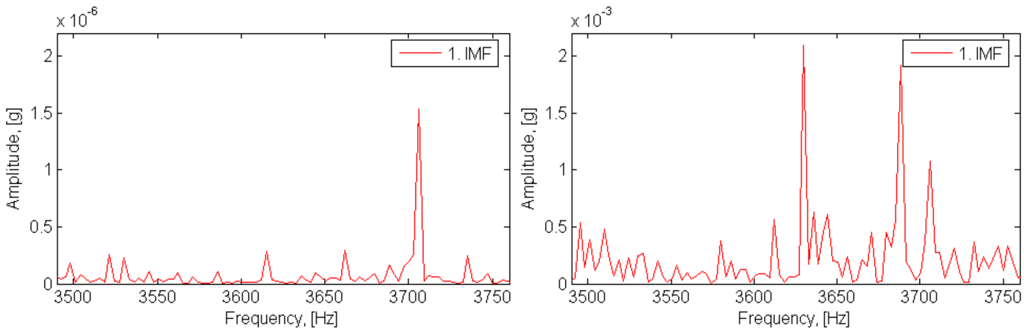

Figure 5.11: The rise of new frequency components and strengthening of the existing ones.

5.4 Analysis of the Vibration Data with Modified Hilbert Huang Transform

Rilling's findings, as presented in Section 3.4.2 and 4.5, defined a limitation on the spectral resolution capability of the EMD process; The EMD could not separate the components of a two-tone signal, to yield two mono-component signals, if the ratio of the frequencies was above 0.67. The rich content of any single IMF estimate actually provides proof for this finding. Also, again in Section 4.5, it was shown that modifications in estimation of the local mean curve and definition of the boundary values presented higher spectral resolution. Presentation of the results obtained by application of modified algorithm again with FFT instead of HT constitutes the goal of this section.

The tests had shown that choosing SD as the termination criteria of sifting increases the number of iterations that may even exceed the number of samples in the data and results in definition of unexpected IMFs. But instead, if the number of siftings is limited to a small number, such as 5, the method produces comparable results. One of the proposals of Huang concerned definition of an *S*-number to limit the number of siftings in order to prevent over-decomposition [3].

The decomposition results of mHHT are depicted in Figure 5.12-5.15. Comparative analysis of Figures 5.8 and 5.13, which present the zoomed in PSDs of the first two IMFs of the initial cycle data extracted by HHT and mHHT, reveals that mHHT provides better estimates of components with higher amplitudes in the PSD.

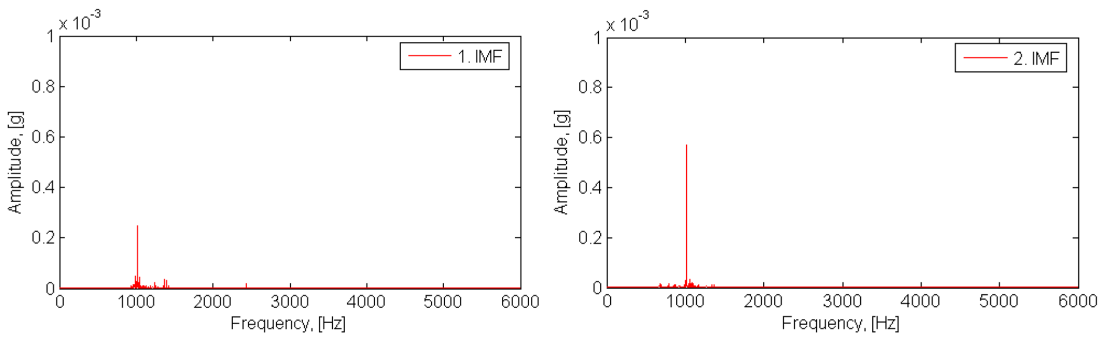

Figure 5.12: PSDs of the first two IMFs of the initial cycle data extracted by mHHT.

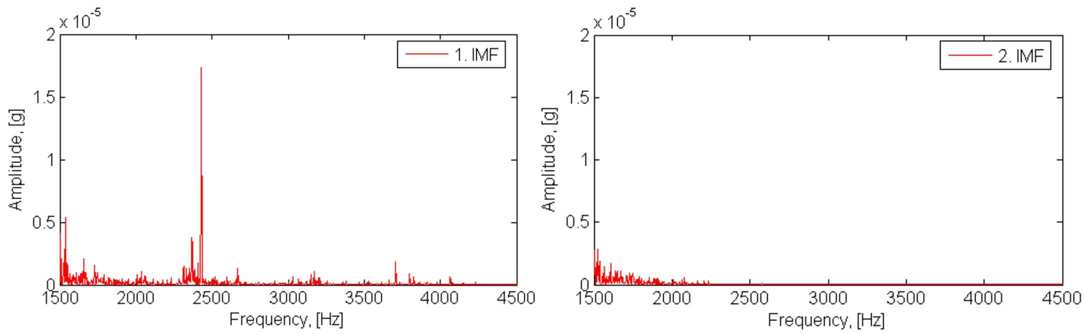

Figure 5.13: Zoomed in values for Figure 5.12.

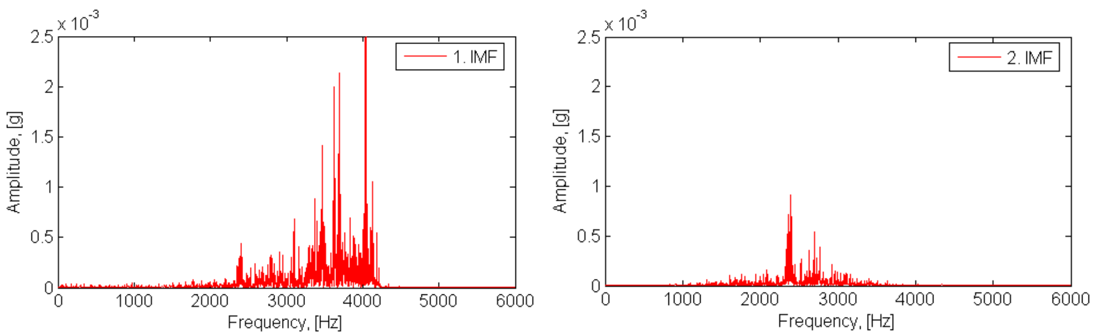

Figure 5.14: PSDs of the 1^{st} and 2^{nd} IMFs of the last cycle data extracted by mHHT.

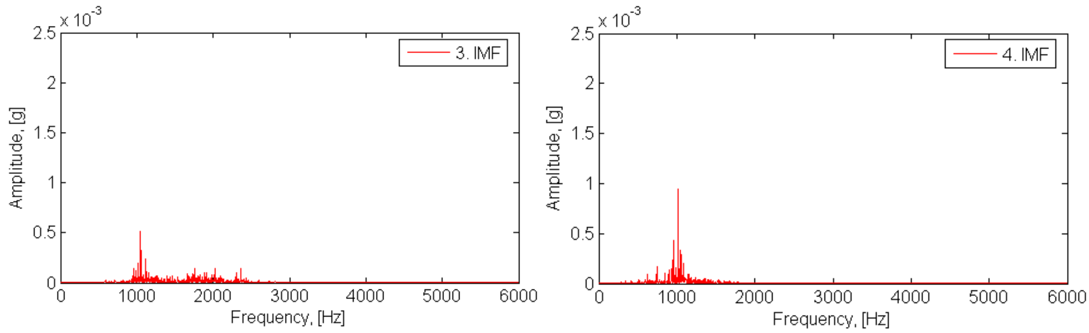

Figure 5.15: PSDs of the 3rd and 4th IMFs of the last cycle data extracted by mHHT.

Rising frequency components in the 2-4 kHz region are retained in the first two IMFs of the last cycle data decomposition and the dominant components of initial cycle are shifted to 3rd and 4th components while they were occupying the first two components previously.

The mHHT provides better resolution even in the low frequency region where EMD is known to be suffering. As an example, f_r is identified as the 14th IMF and almost as a single component.

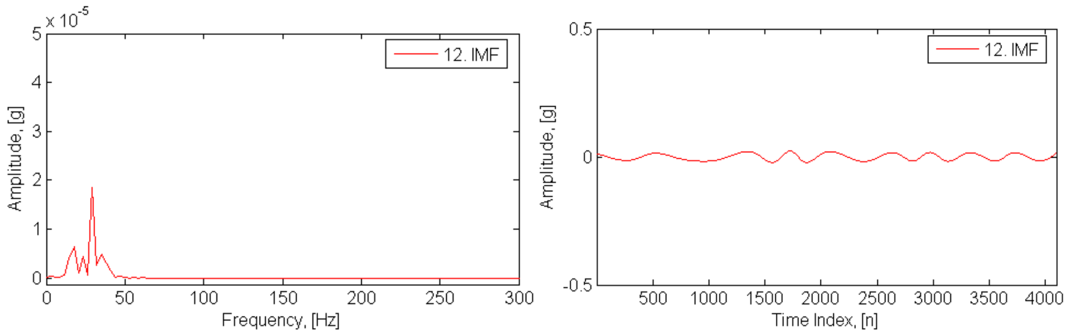

Figure 5.16: f_r in the initial cycle.

The amplification of f_r component due to rotational imbalances imposed by aging process can be detected from comparative analysis of Figure 5.16 and 5.17.

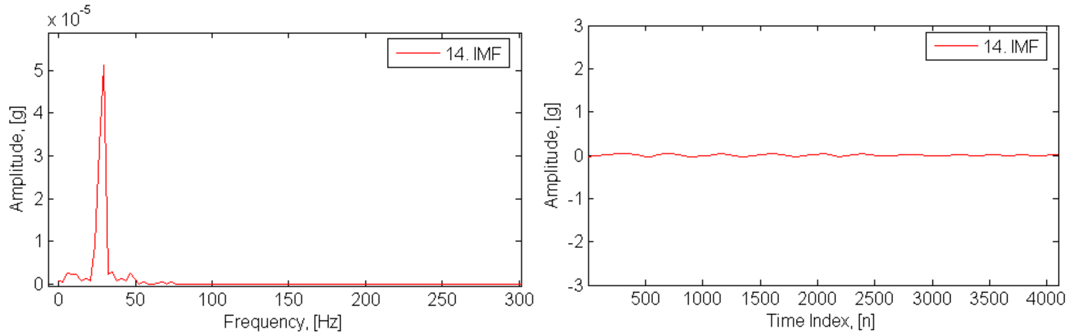

Figure 5.17: f_r in the last cycle.

10th and 11th IMFs of HHT are the most suitable candidates to correspond the same frequency region. When the IMFs are examined in the time-domain, it may be concluded that these content is generated due to inward propagation of the end effect without having a physical meaning.

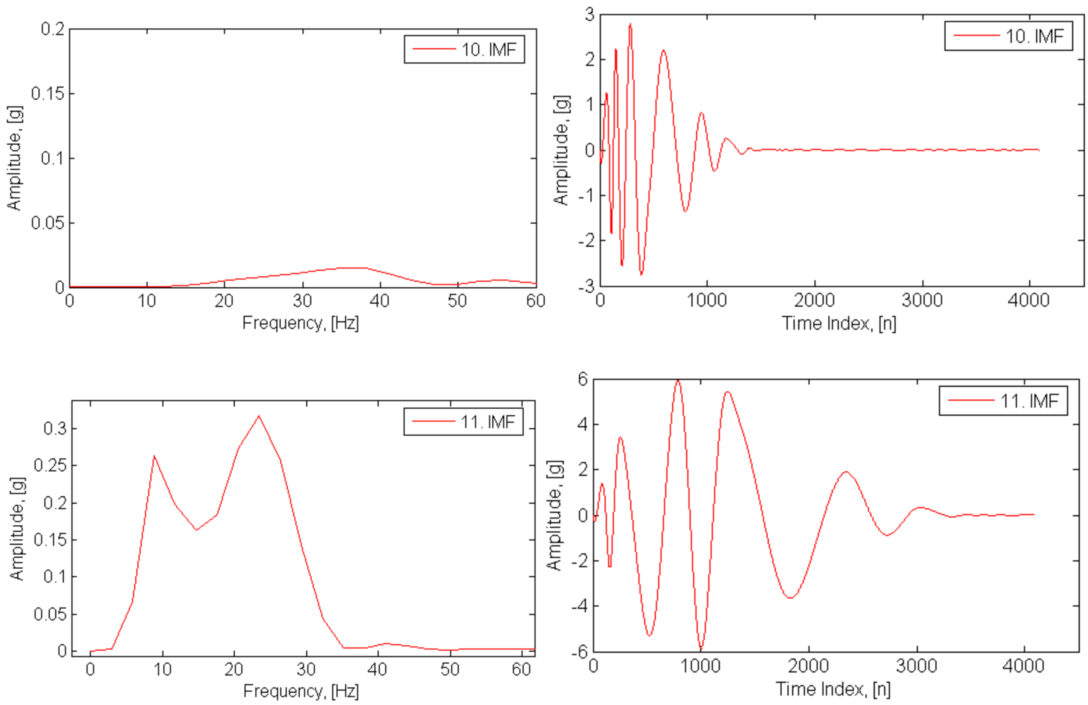

Figure 5.18: HHT components corresponding to 15-55 Hz region.

5.5 On the Computational Performance of the Methods

HHT has already been evaluated as a computationally efficient method [31]. Total numbers of siftings for identification of IMFs with the initial and aged cycle data are given in Table 5.2. It is obvious that mHHT has better computational efficiency.

- **Table 5.2:** Total number of siftings for identification of IMFs.

Method	Number of Siftings at	
	Initial Cycle	Aged Cycle
HHT	3798	8970
mHHT	100	115

6. CONCLUSION AND RECOMMENDATIONS

In this study, a modified HHT (mHHT) algorithm is proposed and the vibration signals acquired during an experimental aging process of an electrical motor is analyzed with CWT, HHT and modified HHT methods.

With the proposed mHHT method, higher spectral resolution beyond the limit imposed with the original algorithm has been achieved and a signal consisting of two sine waves with frequency ratio higher than 0.67 is resolved into its components. In the spectral resolution context, the achievement of IMF identification in the low frequency region should also be noted. In addition, the number of siftings per extraction of a single IMF is remarkably reduced and in return the overall computational time consumption is significantly reduced. Also, the probability of superficial effects due to spline fitting through extrema is reduced in proportion.

Despite it has been known relatively longer, the application of CWT yielded original results showing that the rare small amplitudes in the initial case of the aging process represents the primitive effects of the bearing damage effects occurred at high frequency region of the spectrum of the faulty or aged case data. This finding, based on the power of the continuous wavelet transform with its zoom capability, was presented as an article referred as [8]. This result is also verified with HHT and mHHT within this study.

In addition to extended analysis studies on the frequency resolution performance of the modified algorithm, the following two improvement proposals that concern the termination criteria and “wavelet packet-like” decomposition with mode selection feature are planned to be completed for the near future and will be presented in forthcoming publications.

When the spectral evolution of the frequency components defined with the modified algorithm was considered during the analyses, it was noticed that:

- At the first iterations of HHT and mHHT, the spectra are almost identical,

- Of the wide-band multi-component signal, the low frequency part gets supressed starting from the second iteration and the amplitude of the high frequency part starts to dominate;
- After reaching a maximum value, the amplitude starts to decrease.

Therefore, this observation can be assessed as a termination criterion for sifting. But the number of frequency components may pose a problem. The component with maximum amplitude may be continuously changing at each iteration. To prevent this and further reduce the number of siftings “wavelet packet-like” decomposition with mode selection feature is realized partially.

REFERENCES

[1] **Url-1** <*http://www.nasa.gov/centers/goddard/news/huang_medal.html*>, Access date: 10.05.2010.

[2] **N.E. Huang, et al.,** 1998. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis. *Proceedings of the Royal Society of London*, A 454, p.903–995.

[3] **N.E. Huang, S.S.P. Shen,** 2005: *Hilbert-Huang Transform and Its Applications*. World Scientific Publishing Co. Pte. Ltd., Singapore.

[4] **N.E. Huang, M.L. Wu, W. Qu, S.R. Long, S.S.P. Shen, J.E. Zhang,** 2003. Applications of Hilbert-Huang Transform to Non-stationary Financial Time Series Analysis. *Applied Stochastic Models In Business and Industry*, **19**, p.245-268.

[5] **S. Şeker, E. Ayaz, B.R. Upadhyaya and A.S. Erbay,** 2000. Analysis of Motor Current and Vibration Signals for Detecting Bearing Damage in Electric Motors. *Maintenance and Reliability Conference Proceedings*, Knoxville, TN, USA, p.29.01-29.14.

[6] **A.S. Erbay,** 1999. Multi Sensor Fusion for Induction Motor Aging Analysis and Fault Diagnosis. *PhD thesis*, The University of Tennessee, Knoxville, TN.

[7] **M.J. Devaney, L. Eren,** 2004: Detecting Motor Bearing Faults. *IEEE Instrumentation & Measurement Magazine*, p.30-50.

[8] **E. Ayaz, A. Öztürk, S. Şeker,** 2009. Fault Detection Based on Continuous Wavelet Transform and Sensor Fusion in Electric Motors. *Compel*, Vol. 28, No. 2, p. 454-470

[9] **L. Eren, A. Karahoca, M.J. Devaney,** 2004. Neural Network Based Motor Bearing Fault Detection. *IMTC - Instrumentation and Measurement Technology Conference*, Como, Italy.

[10] **S.H. Ghafari**, 2007. A Fault Diagnosis System for Rotary Machinery Supported by Rolling Element Bearing, *PhD Thesis*, University of Waterloo, Waterloo, Ontario, Canada.

[11] **J. Mathew, R.J. Alfredson,** 1983. The condition monitoring of rolling element bearings using vibration analysis. *American Society of Mechanical Engineers - Winter Annual Meeting*, Boston, MA; United States.

[12] **S. Şeker, E. Ayaz,** 2003: Feature Extraction Related to Bearing Damage in Electric Motors by Wavelet Analysis. *Journal of The Franklin Institute*, **340**, p.125-134.

[13] **S.A. Mcinerny, Y. Dai,** 2003. Basic Vibration Signal Processing for Bearing Fault Detection. *IEEE Transactions on Education*. Vol. **46**, No 1, p.149-156.

[14] **J.S. Bendat, A.G. Piersol,** 1986. *Random Data: Analysis and Measurement Procedures*, John Wiley & Sons Inc., USA.

[15] **A.Y. Ayenu-Prah and N.O. Attoh-Okine,** 2009. Comparative study of Hilbert–Huang Transform, Fourier Transform and Wavelet Transform in pavement profile analysis. *Vehicle System Dynamics*, Vol. **47**, No. 4, p.437–456.

[16] **M. Chendeb, M. Khalil, J. Duchêne,** 2006. The Use of Wavelet Packets for Event Detection. *Journal of Signal Processing - Special Section: Multimodal Human-Computer Interfaces*, Vol.**86**, Issue 12. Elsevier North-Holland, Inc. Amsterdam, The Netherlands.

[17] **V. Vatchev,** 2002. The Analysis of the Empirical Mode Decomposition Method. *University of Southern California Lecture Notes*, 2002.

[18] **N. S. Majumdar**, 2007. Signal Analysis with Huang Hilbert Transforms, *PhD Thesis,* George Mason University, Fairfax, VA, USA.

[19] **Q. Gao, C. Duan, H. Fan, Q. Meng ,** 2008. Rotating Machine Fault Diagnosis Using Empirical Mode Decomposition. *Mechanical Systems and Signal Processing*. **22**, p.1072–1081.

[20] **M. Johansson,** 1999. The Hilbert Transform, *Master's Thesis*, Vaxjo University.

[21] **D.G. Long,** 2004. Comments on Hilbert Transform Based Signal Analysis. *MERS Technical Report*, **# MERS 04-001**, Microwave Earth Remote Sensing (MERS) Laboratory, Brigham Young University, Utah, USA.

[22] **J.M.H. Peters,** 1995. A beginner's guide to the Hilbert Transform', International *Journal of Mathematical Education in Science and Technology*, Vol. **26**, No. 1, p.89-106

[23] **G. Rilling, P. Flandrin,** 2008. One or Two Frequencies? The Empirical Mode Decomposition Answers. *IEEE Trans. On Signal Processing*, Vol. **56**, Issue 1, p.85-95.

[24] **Z.H. Wu,N.E. Huang,** 2009. Ensemble Empirical Mode Decomposition: a noise assisted data analysis method. *Advances in Adaptive Data Analysis*, Vol. **1**, p.1–41

[25] **H. Li, L. Yang, D. Huang,** 2005. The study of the intermittency test filtering character of Hilbert–Huang transform. *Mathematics and Computers in Simulation*, Vol. **70**, Issue 1, p.22-32.

[26] **D. Duanhui, W. Qiusheng,** 2006. An improved empirical mode decomposition based on combining extrapolating extrema with mirror extension. *Proc. Of SPIE*, Vol. **6357**, p. 63570X-1:6357X-6

[27] **F.Wu, L. Qu,** 2008. An improved method for restraining the end effect in empirical mode decomposition and its applications to the fault diagnosis of large rotating machinery. *Journal of Sound and Vibration*, **314**, p. 586–602.

[28] **K.T. Coughlin, K.K. Tung,** 2004. 11-Year solar cycle in the stratosphere extracted by the empirical mode decomposition method, *Advances in Space Research* **34**, p.323–329.

[29] **G. Rilling, P. Flandrin, P. Goncalves,** 2003. On empirical mode decomposition and its algorithms, *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, **NSIP-03**, p.1–5.

[30] **W.X. Yang,** 2008. Interpretation of mechanical signals using an improved Hilbert–Huang transform. *Mechanical Systems and Signal Processing*, **22**, p.1061–1071.

[31] **Z.K. Peng, P.W. Tse, F.L. Chu,** 2005. A comparison study of improved Hilbert–Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing. *Mechanical Systems and Signal Processing*, **19**, p.974–988

[32] **G. Rilling, P. Flandrin,** 2009. Sampling effects on the Empirical Mode Decomposition. *Advances in Adaptive Data Analysis (AADA)*, Vol.**1**, No. 1, p.43-59.

[33] **D. Donelly,** 2006. The Fast Fourier and Hilbert-Huang Transforms: A Comparison. *International Journal of Computers, Communications & Control*, Vol. **I**, No. 4, p.45-52

[34] **P. Flandrin, P. Gonçalvès,** 2004. Empirical Mode Decompositions As Data-Driven Wavelet-Like Expansions. *International Journal of Wavelets, Multiresolution and Information Processing*, Vol. **2**, No. 4, p.1–20.

[35] **L. Eren, M.J. Devaney,** 2004: Bearing Damage Detection via Wavelet Packet Decomposition of the Stator Current. *IEEE Transactions on Instrumentation and Measurement Magazine*, Vol. **53**, No. 2, p.431-436.

[36] **V.K. Rai, A.K. Mohanty,** 2006. Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform. *Mechanical Systems and Signal Processing*, **21**, p.2607–2615.

[37] **Z. Peng, P. Tse, F. Chu,** 2005. An improved Hilbert–Huang transform and its application in vibration signal analysis, *Journal of Sound and Vibration*, **286**, p.187–205.

[38] **T.K. Correa, A. Kareem,** 2007. Performance of Wavelet Transform and Empirical Mode Decomposition in Extracting Signals Embedded in Noise. *Journal of Engineering Mechanics*, **133:7**, p.849-852.

APPENDICES

APPENDIX A.1 : Analysis Results of HHT – Initial Cycle Data
APPENDIX A.2 : Analysis Results of Modified HHT – Initial Cycle Data
APPENDIX A.3 : Analysis Results of HHT – Aged Cycle Data
APPENDIX A.4 : Analysis Results of Modified HHT – Aged Cycle Data

APPENDIX A.1

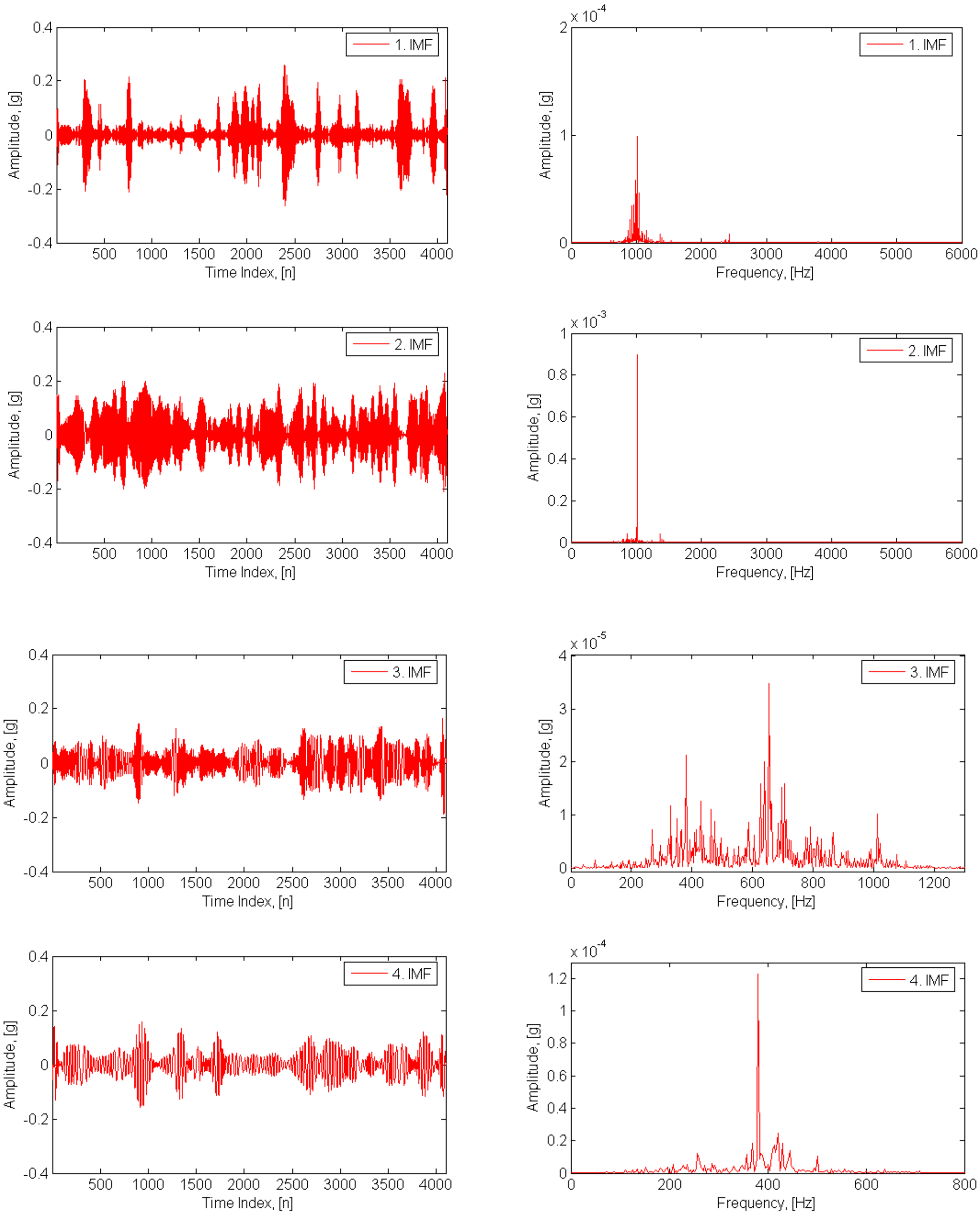

Figure A1.1: Analysis Results of HHT – Initial Cycle Data.

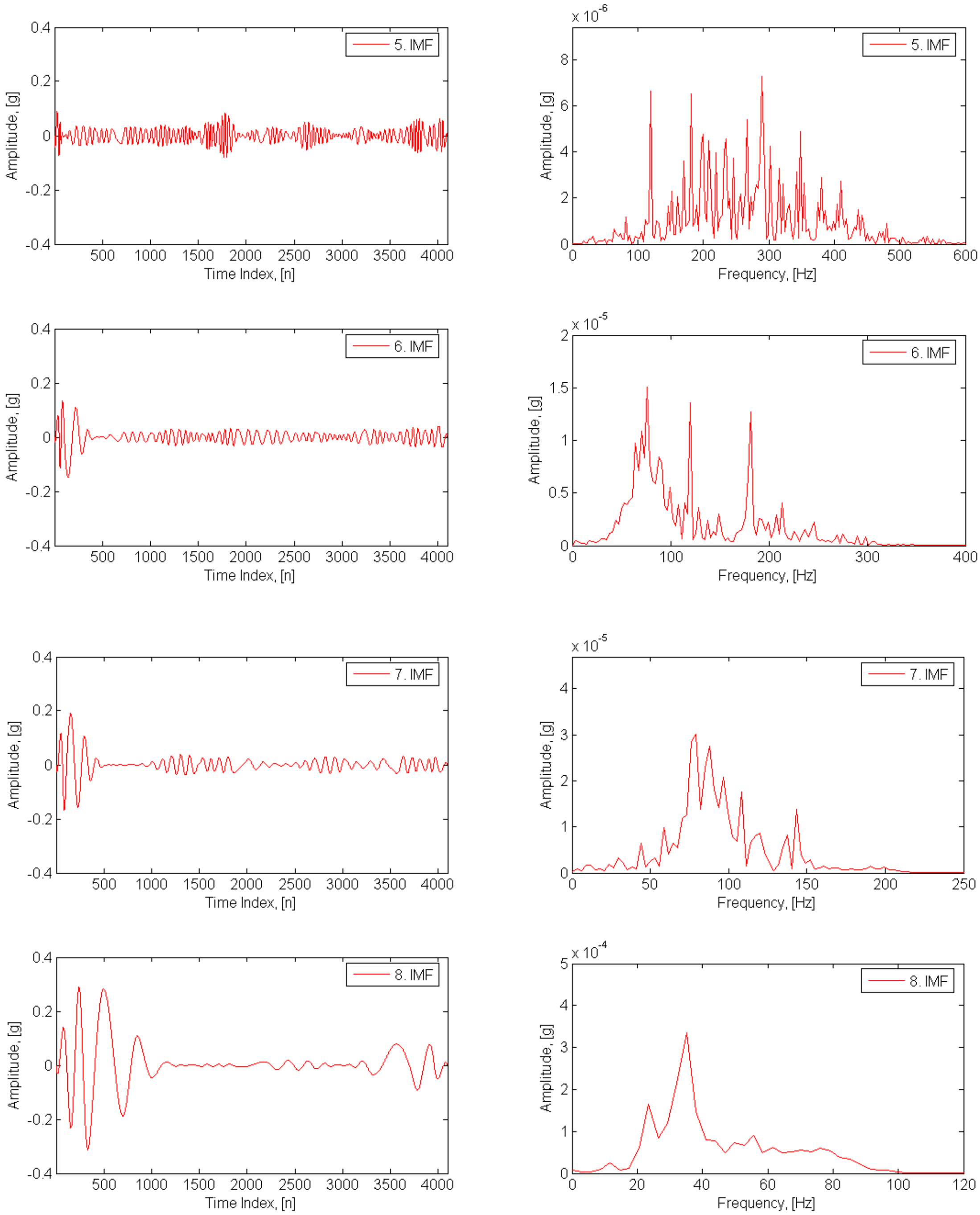

Figure A1.1 (Continued): Analysis Results of HHT – Initial Cycle Data.

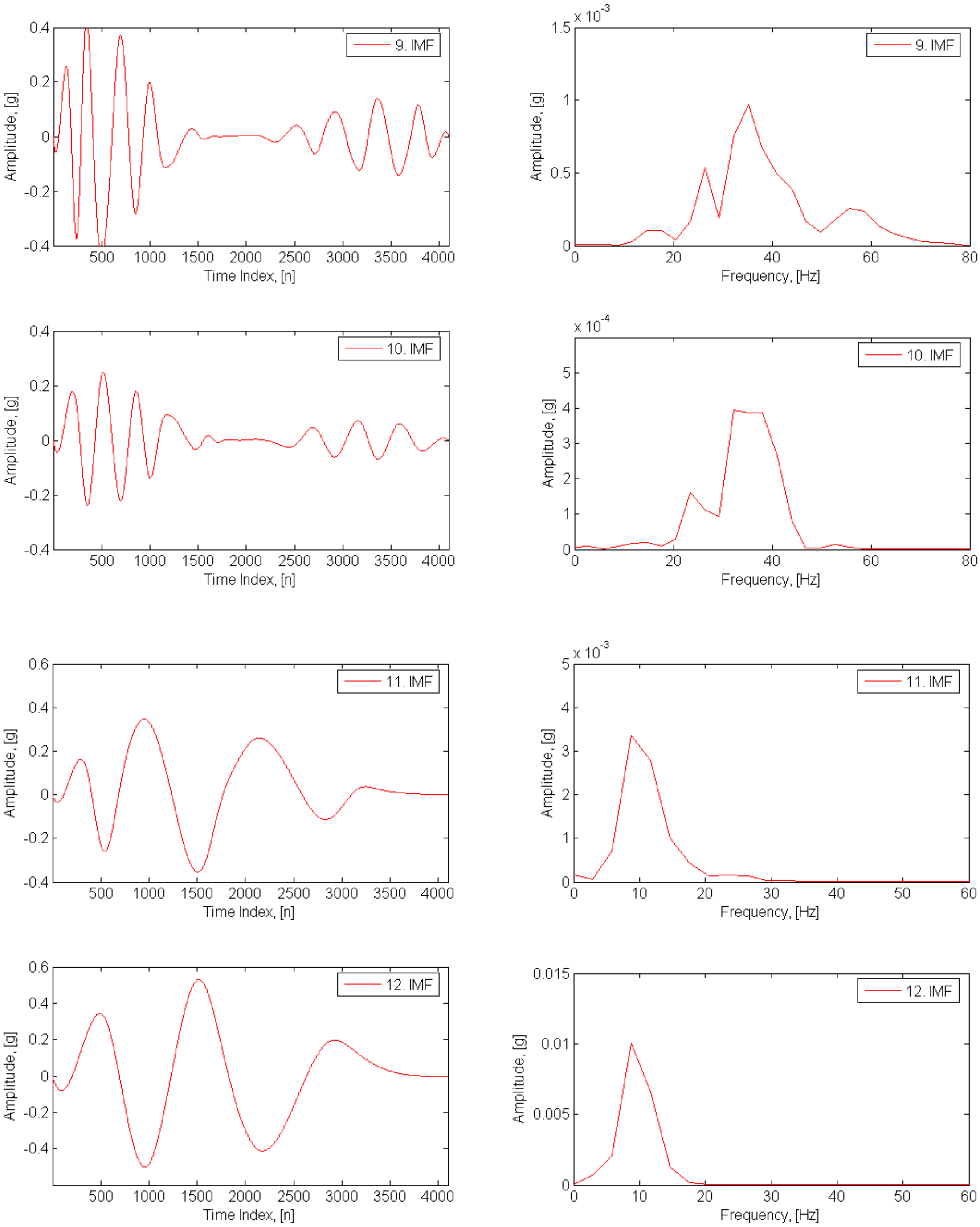

Figure A1.1 (Continued): Analysis Results of HHT – Initial Cycle Data.

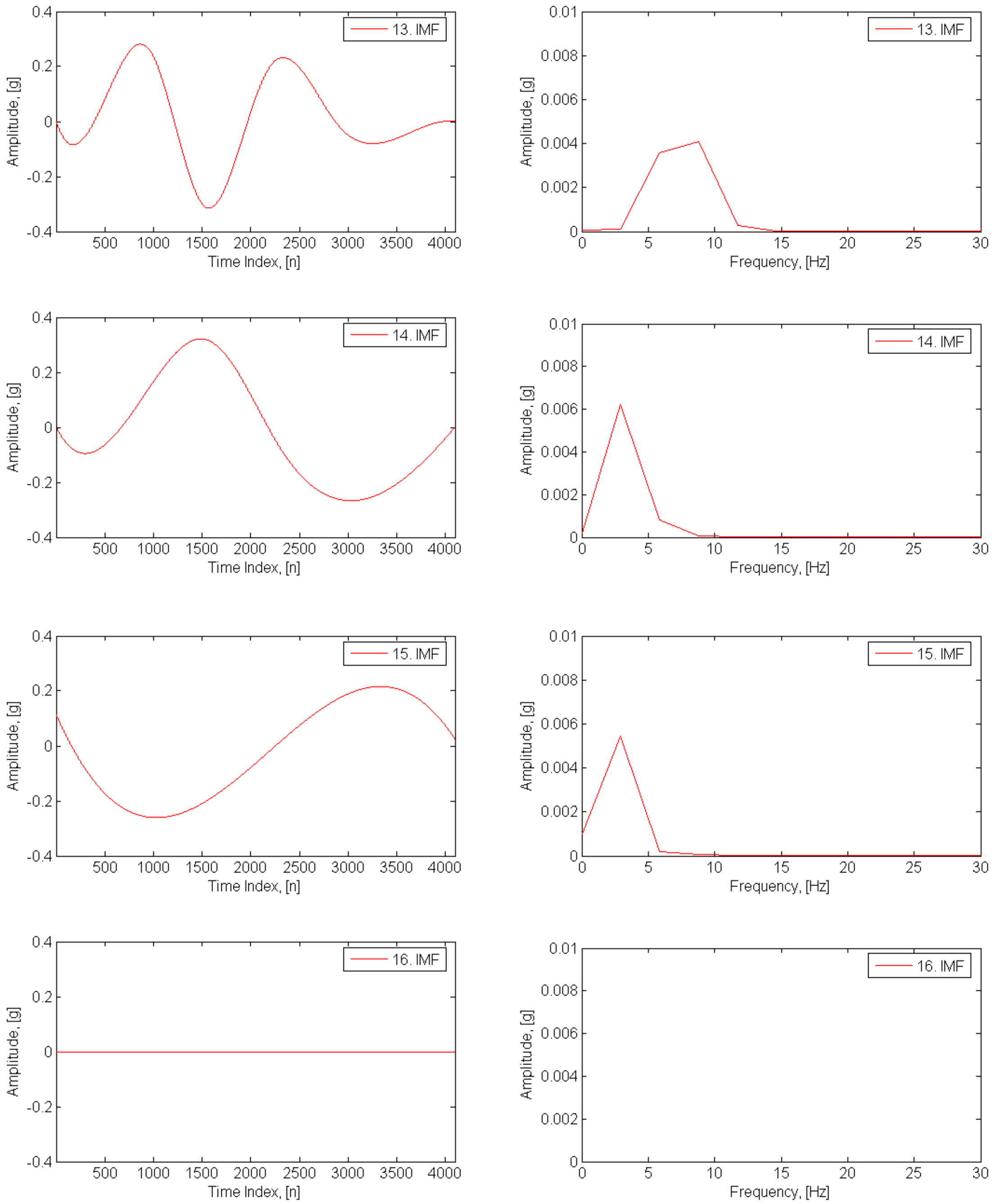

Figure A1.1 (Continued): Analysis Results of HHT – Initial Cycle Data.

APPENDIX A.2

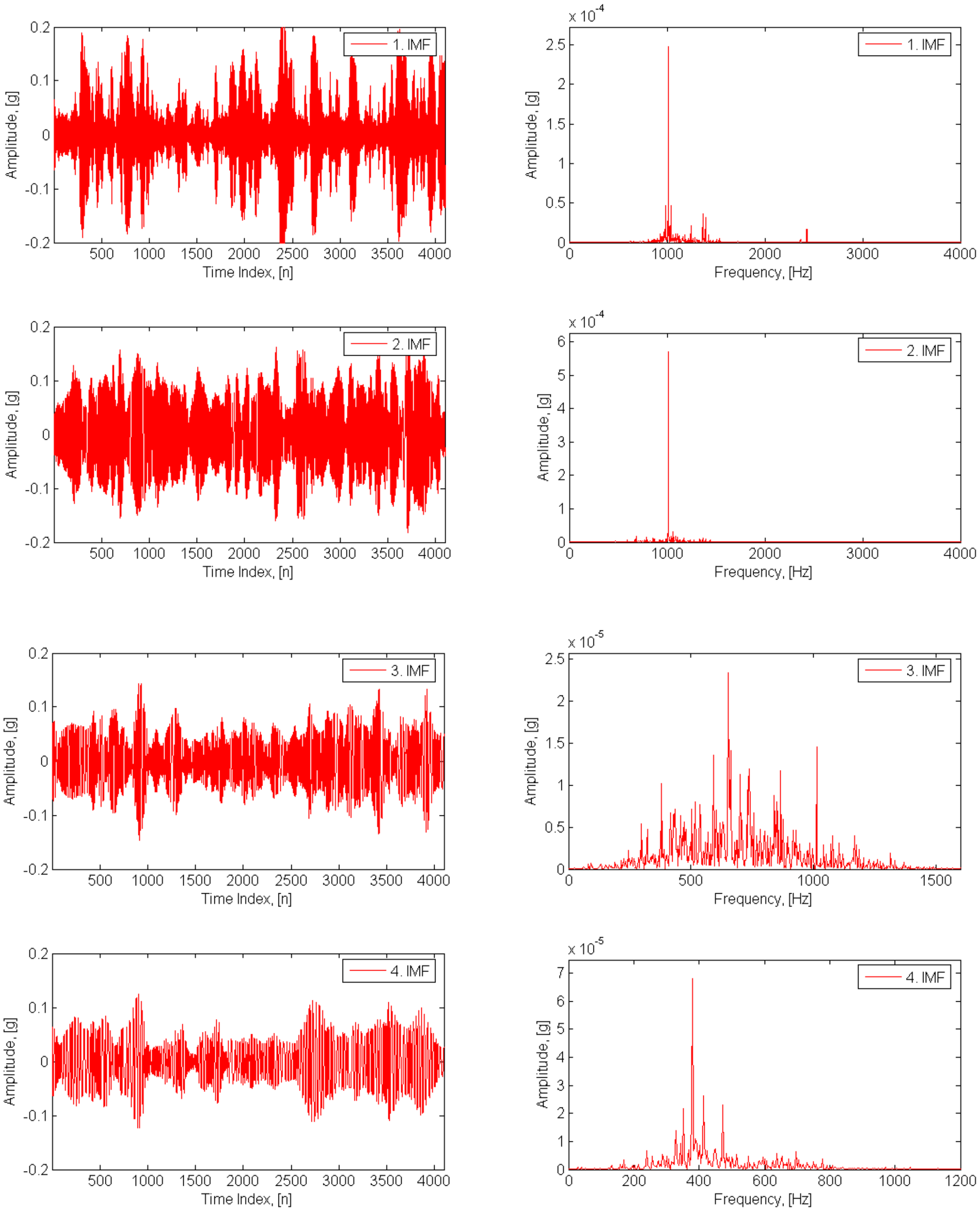

Figure A2.1: Analysis Results of Modified HHT – Initial Cycle Data.

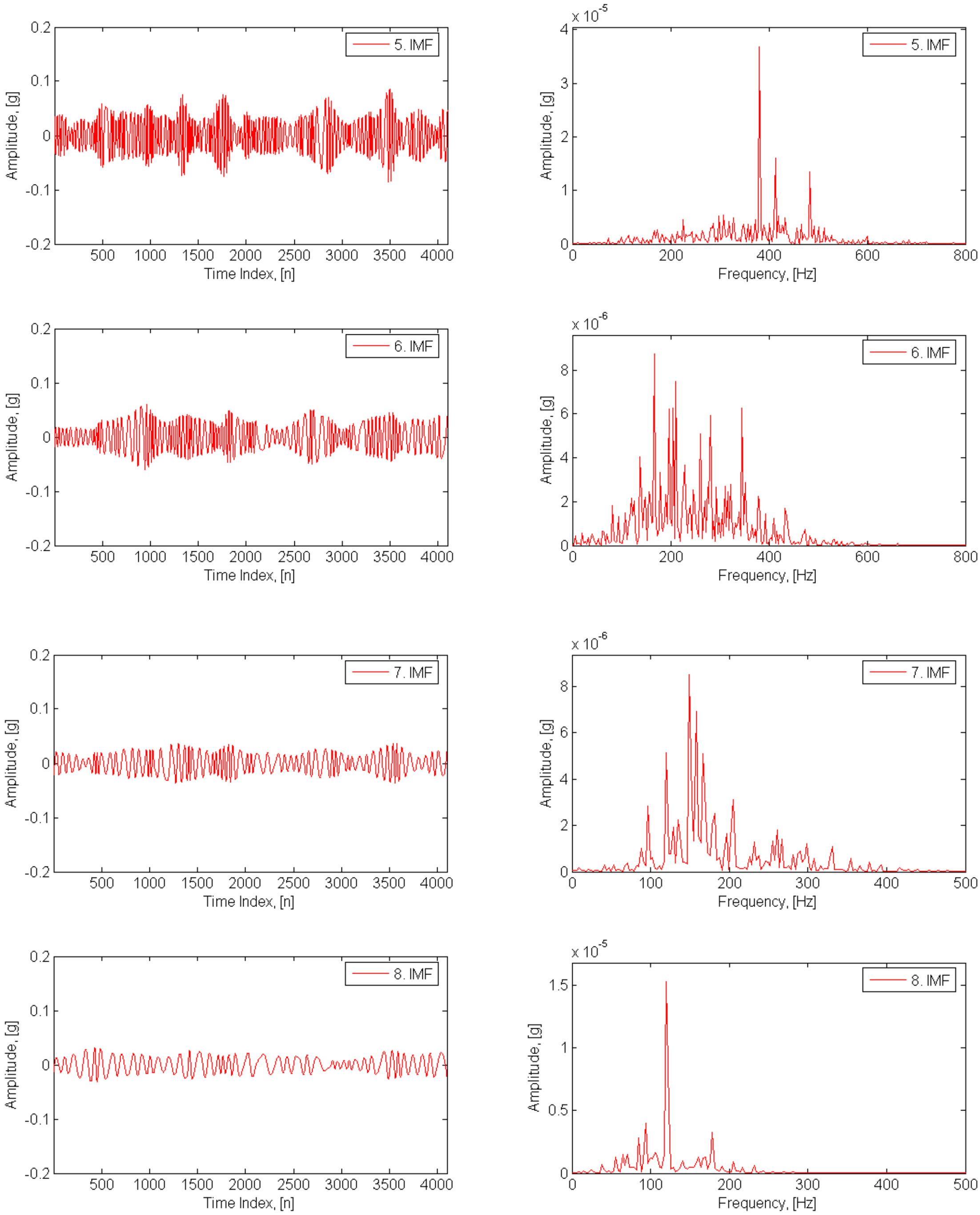

Figure A2.1 (Continued): Analysis Results of Modified HHT – Initial Cycle Data.

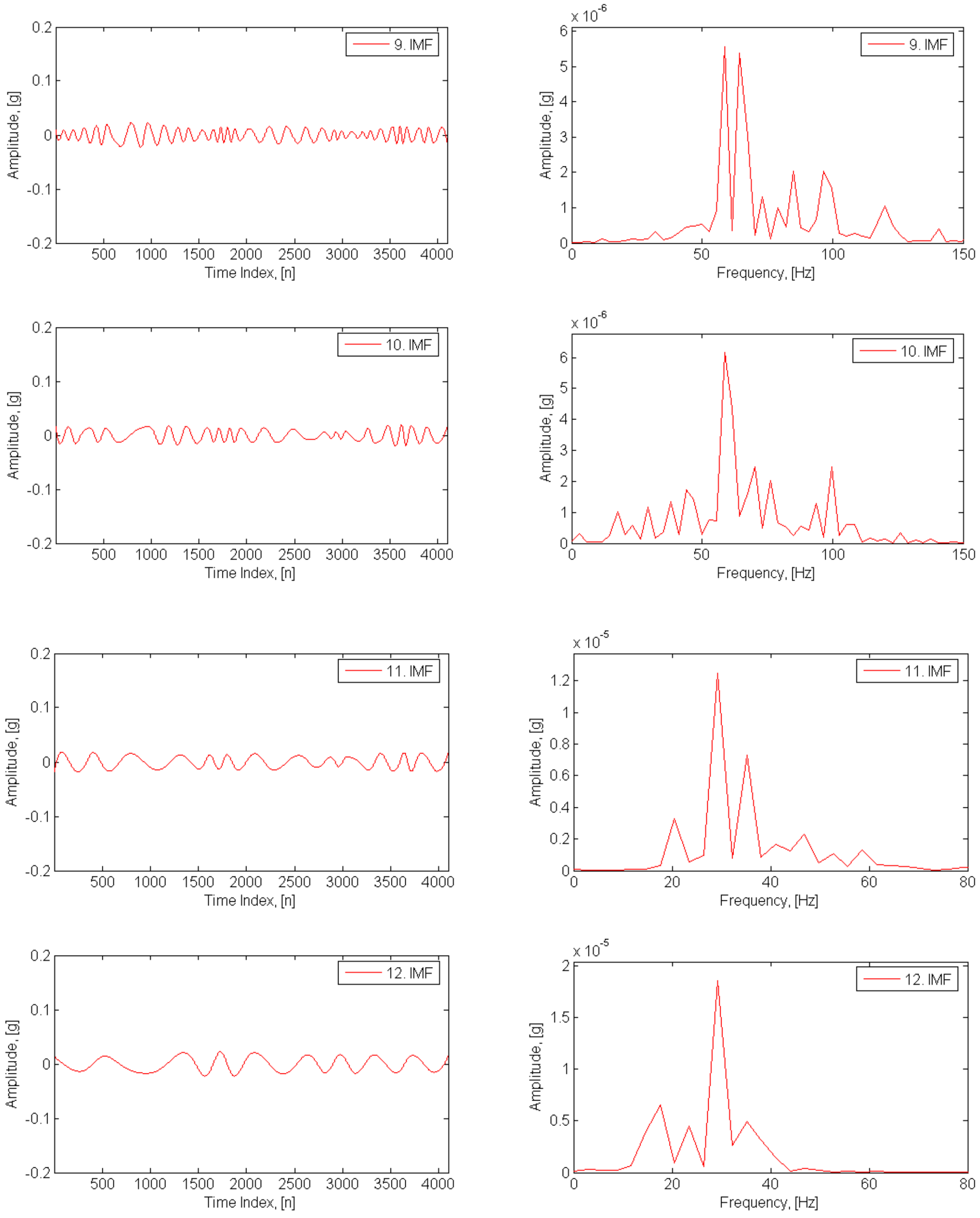

Figure A2.1 (Continued): Analysis Results of Modified HHT – Initial Cycle Data.

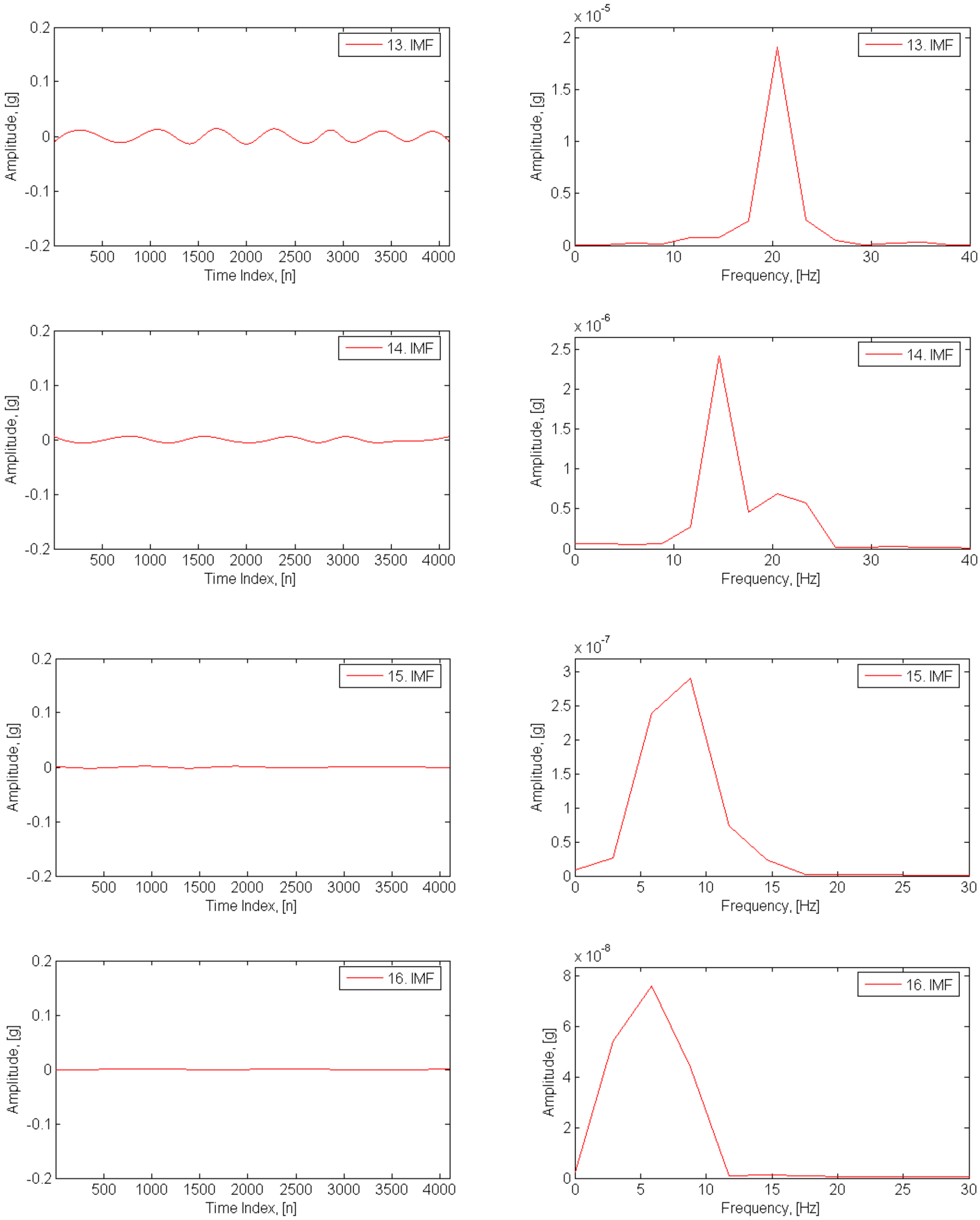

Figure A2.1 (Continued): Analysis Results of Modified HHT – Initial Cycle Data.

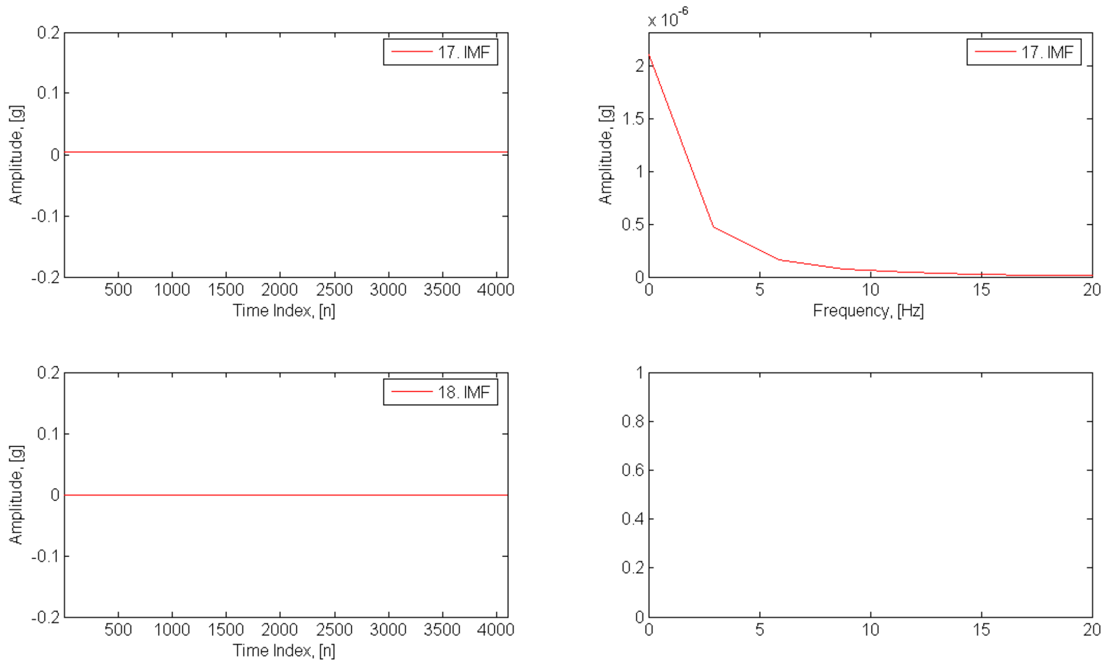

Figure A2.1 (Continued): Analysis Results of Modified HHT – Initial Cycle Data.

APPENDIX A.3

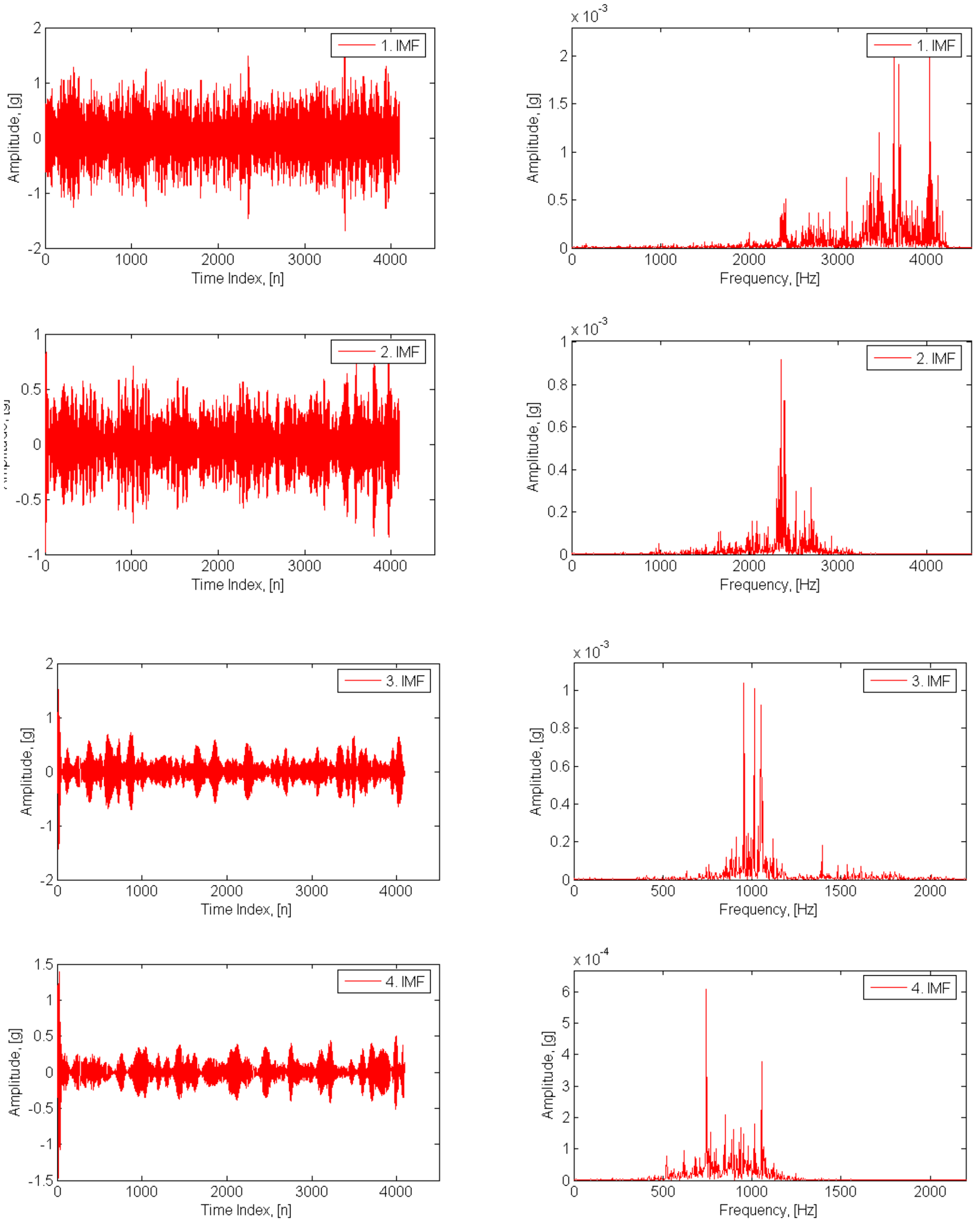

Figure A3.1: Analysis Results of HHT – Aged Cycle Data.

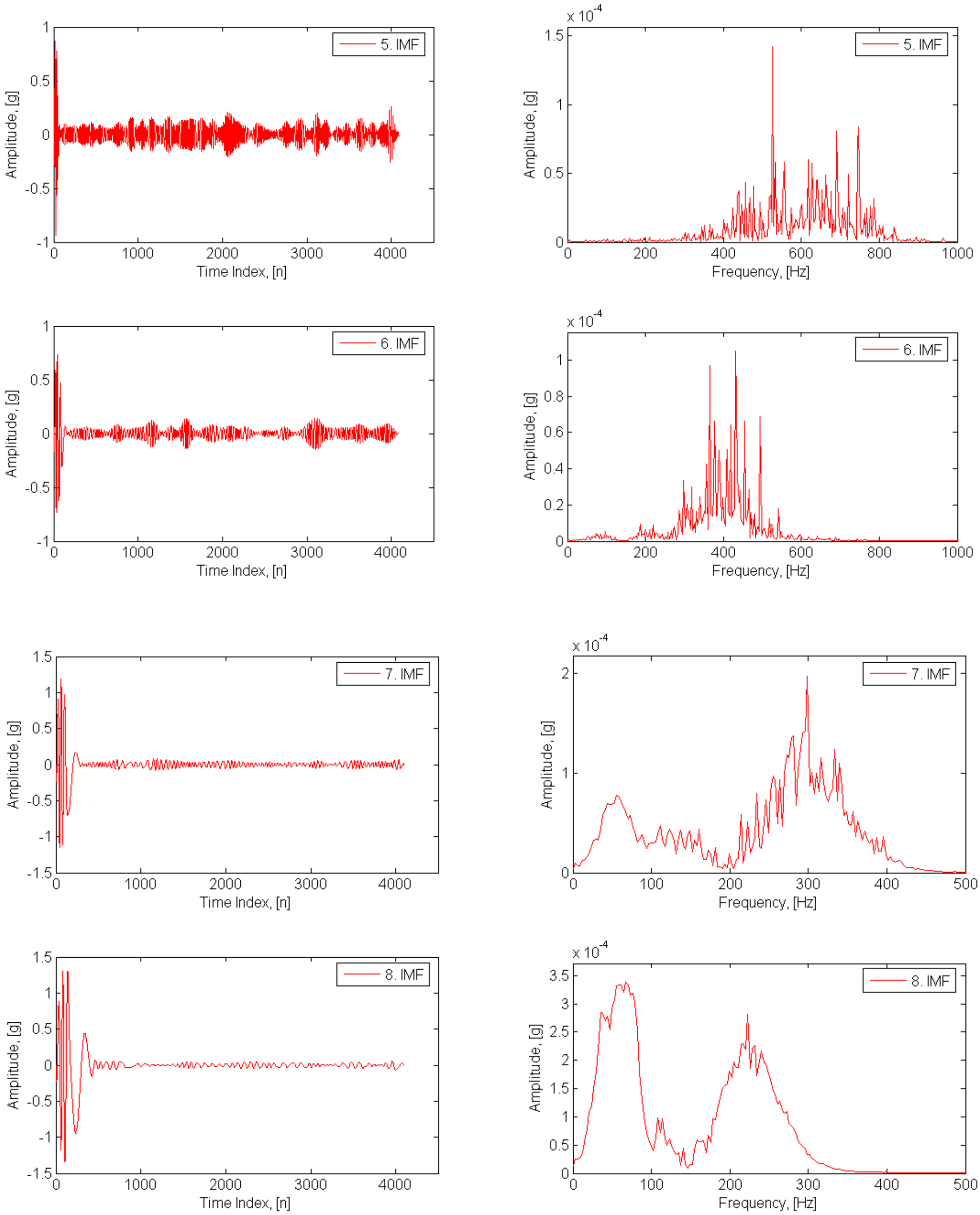

Figure A3.1 (Continued): Analysis Results of HHT – Aged Cycle Data.

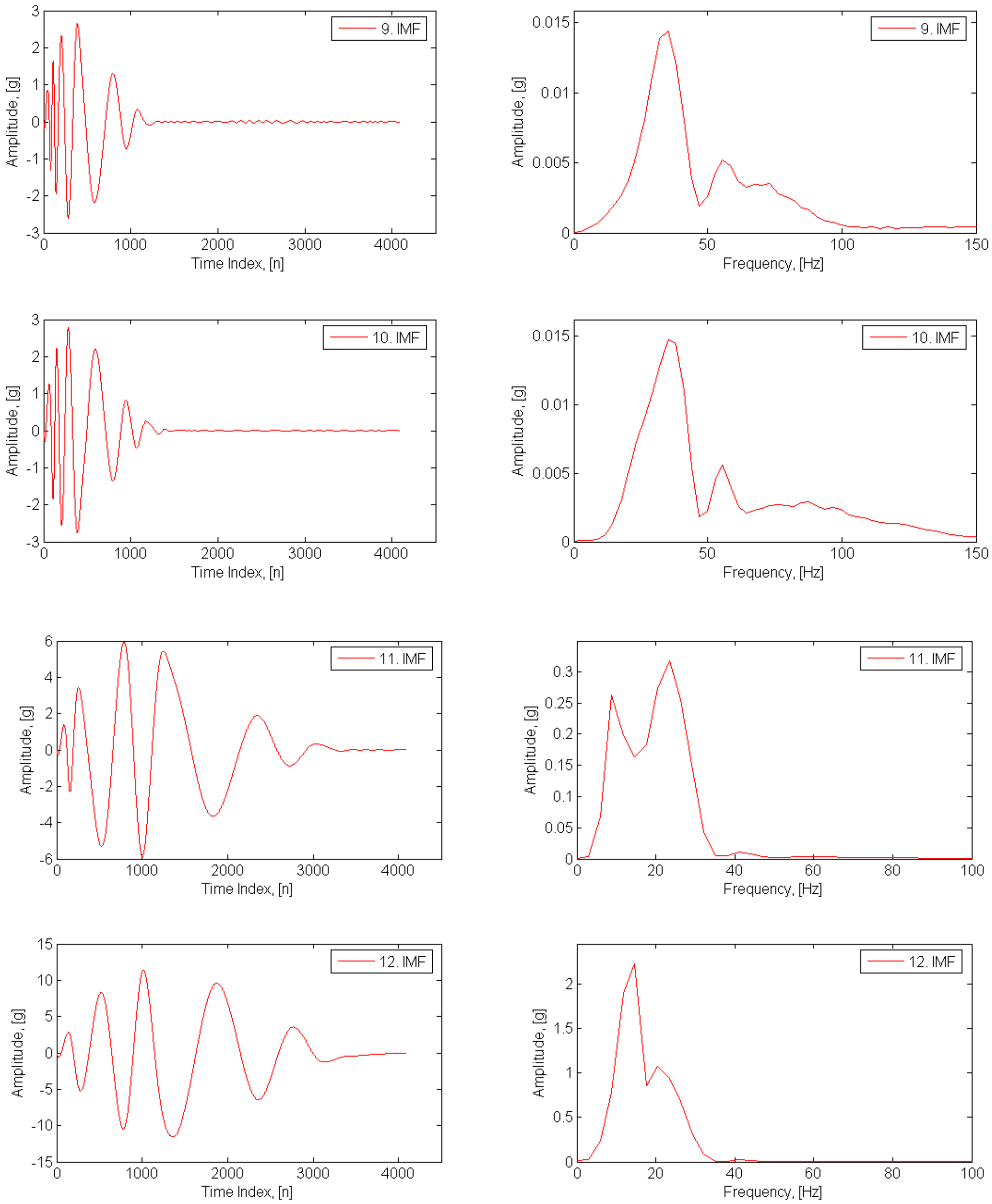

Figure A3.1 (Continued): Analysis Results of HHT – Aged Cycle Data.

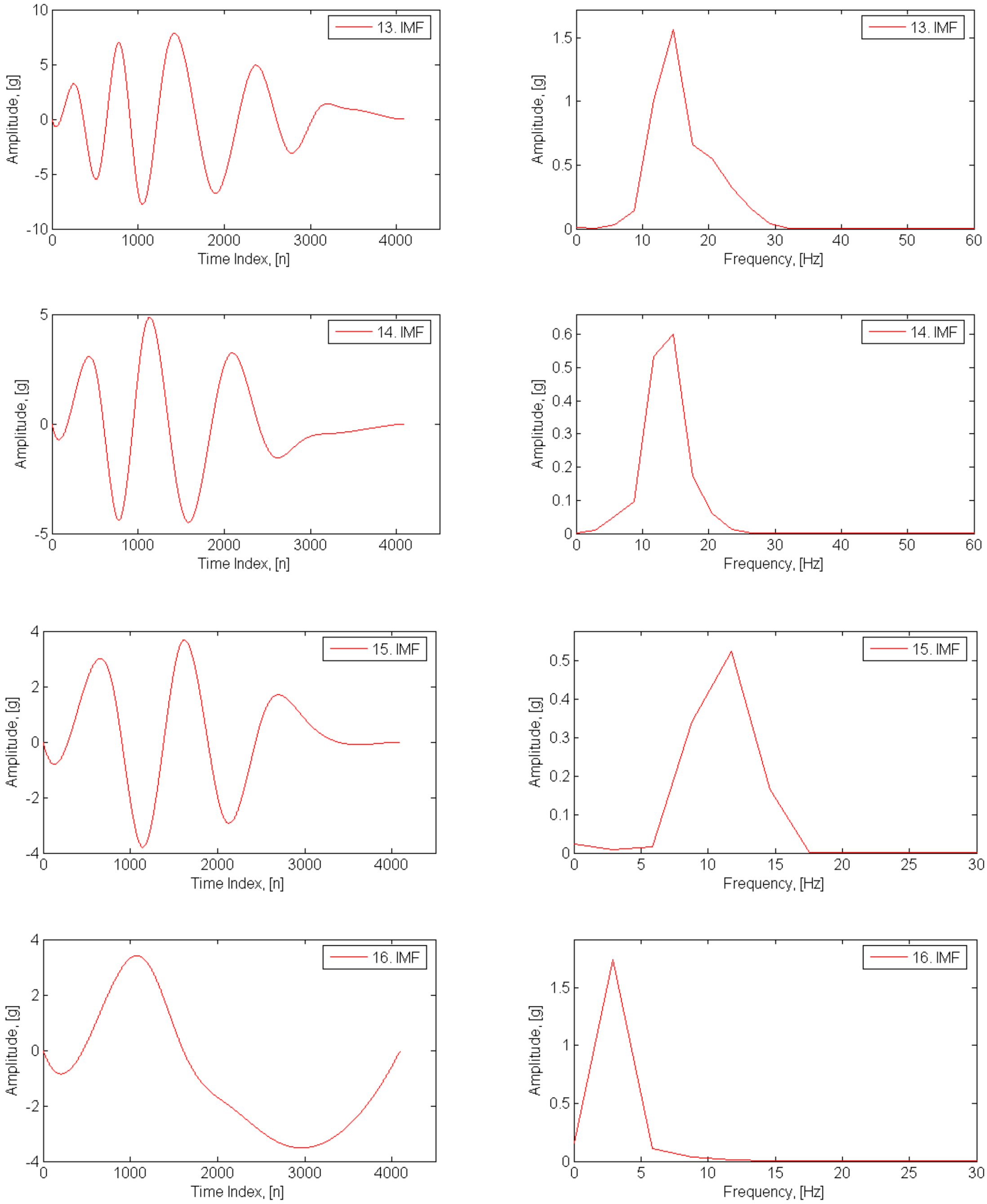

Figure A3.1 (Continued): Analysis Results of HHT – Aged Cycle Data.

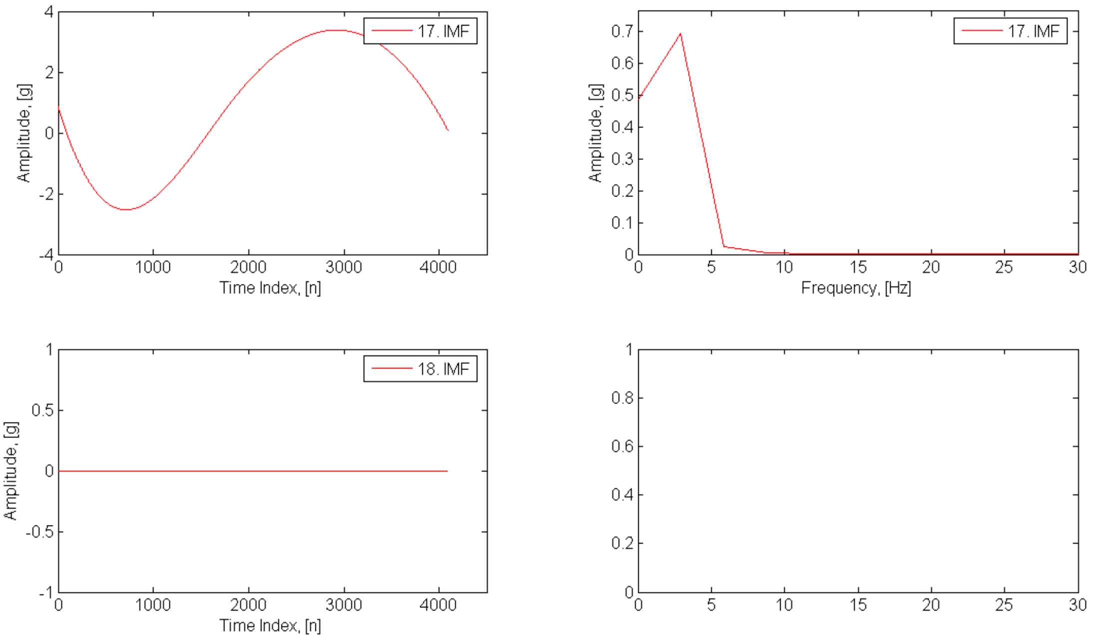

Figure A3.1 (Continued): Analysis Results of HHT – Aged Cycle Data.

APPENDIX A.4

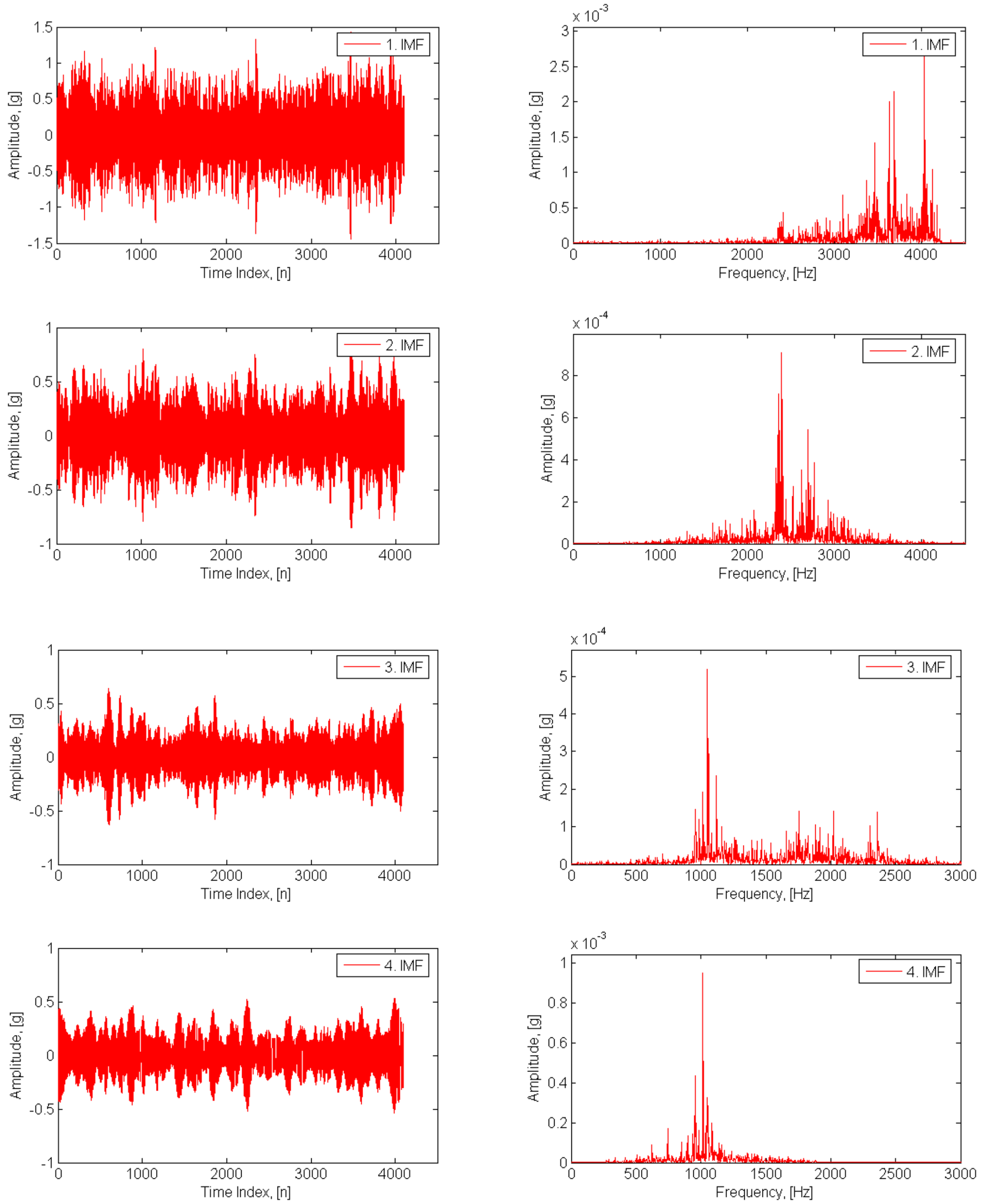

Figure A4.1: Analysis Results of Modified HHT – Aged Cycle Data.

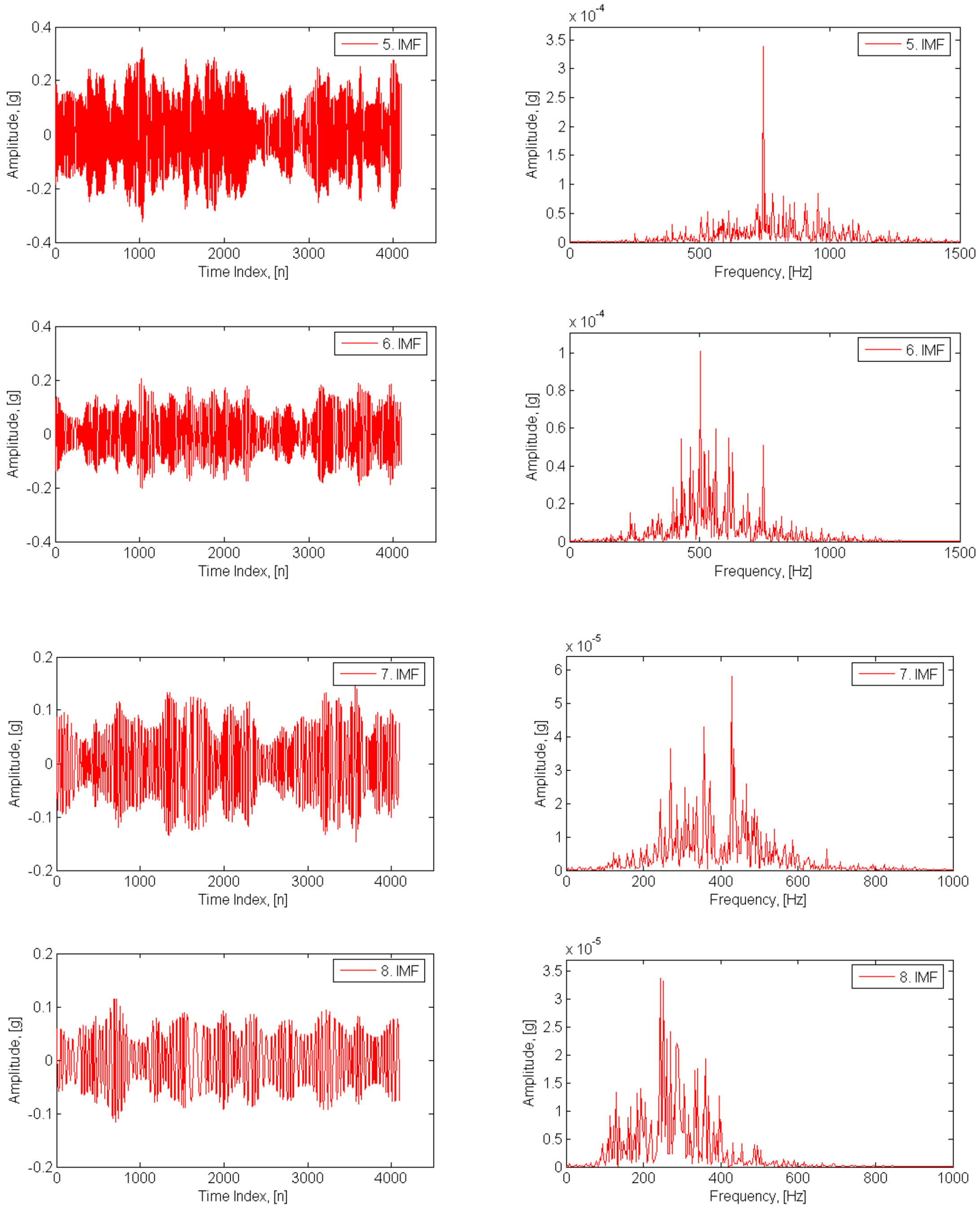

Figure A4.1 (Continued): Analysis Results of Modified HHT – Aged Cycle Data.

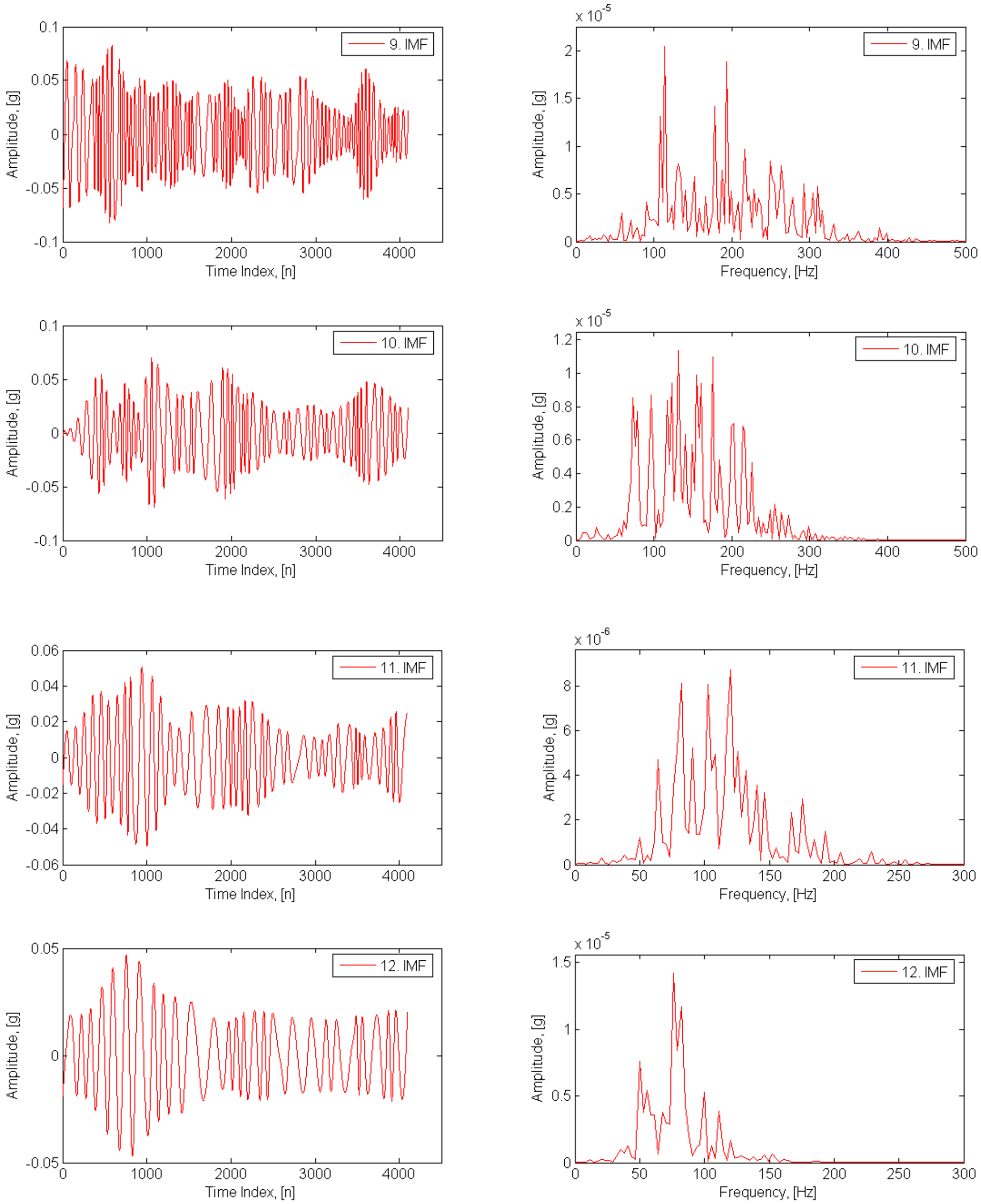

Figure A4.1 (Continued): Analysis Results of Modified HHT – Aged Cycle Data.

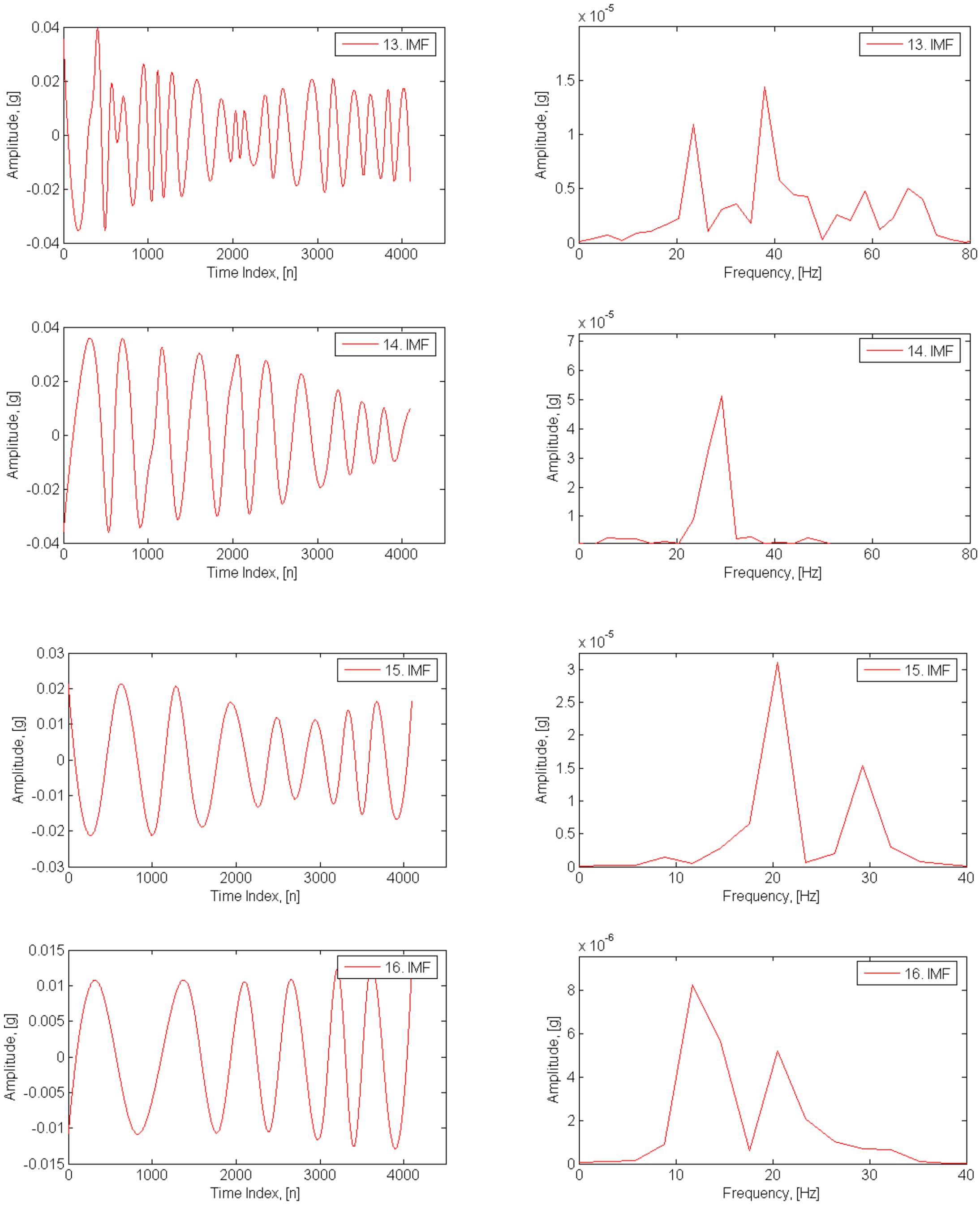

Figure A4.1 (Continued): Analysis Results of Modified HHT – Aged Cycle Data.

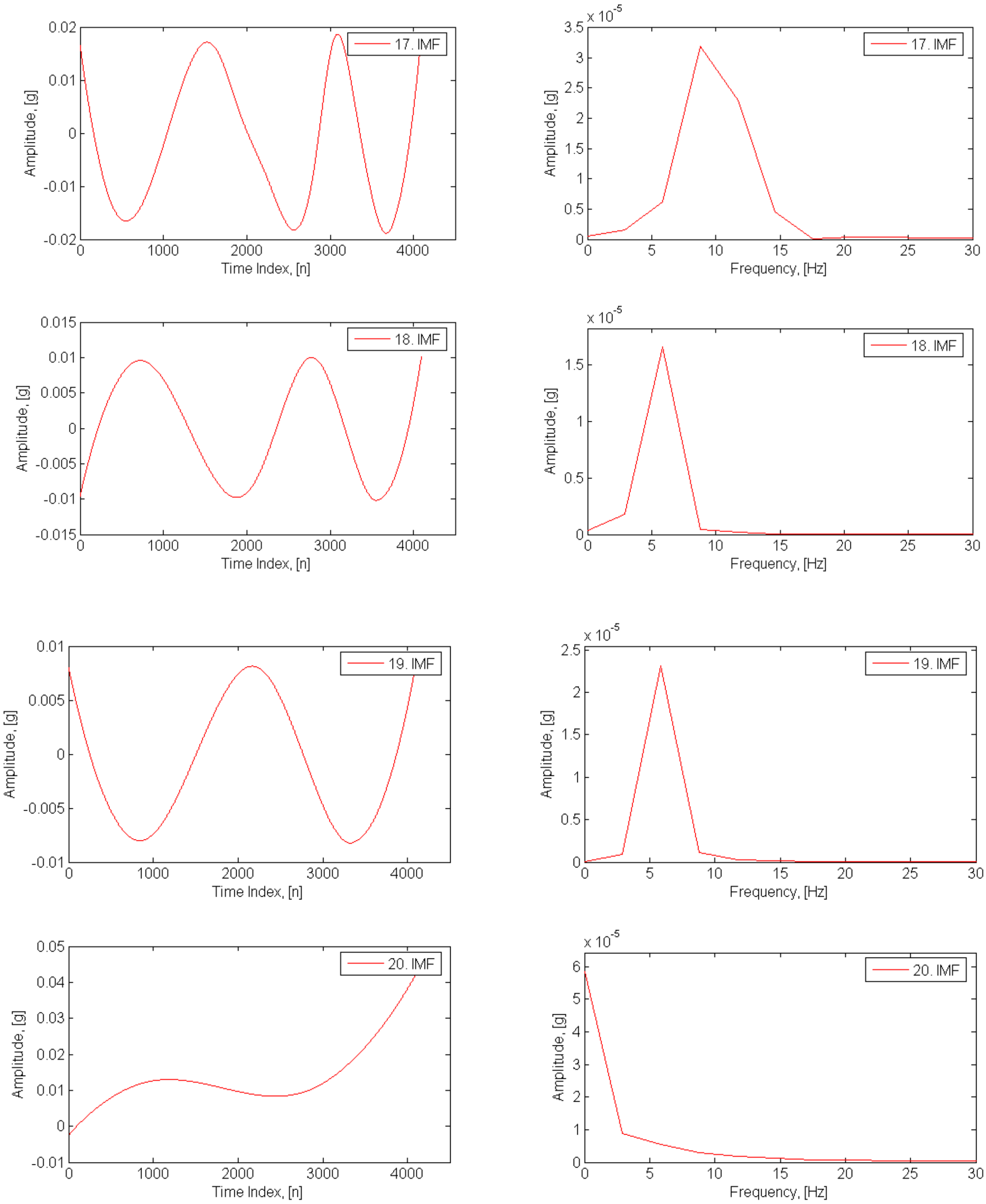

Figure A4.1 (Continued): Analysis Results of Modified HHT – Aged Cycle Data.

CURRICULUM VITA

Candidate's full name: Ahmet Öztürk

Place and date of birth: Bulgaria, 27.12.1974

Permanent Address: Uğur Mumcu Mah. 2305 Sk. 14/2, 34265, Sultangazi, İstanbul, Turkey

Universities and Colleges attended: Istanbul Technical University

Publications:

- **A. Öztürk,** M. Afacan, 2011. "On the Evolution of Safety Concept and Safety Management System of Turkish Airlines", *VI. National Aeronautical and Aerospace Engineering Congress*, Eskisehir, Turkey.
- **A. Öztürk,** S. Şeker, 2010. "On The Frequency Resolution of Improved Empirical Mode Decomposition Method", International Review of Electrical Engineering, Vol. 5 N. 4-- Part B, p.1798-1805.
- E. Ayaz, **A. Öztürk,** S. Şeker, 2009. Fault Detection Based on Continuous Wavelet Transform and Sensor Fusion in Electric Motors. Compel, Vol. 28, No. 2, p. 454-470.
- K. Cığızoğlu, **A. Öztürk,** 2007. "Artificial Neural Network Models In Rainfall-Runoff Modeling Of Turkish Rivers", *International River Basin Management Congress*, Antalya, Turkey.
- K. Cığızoğlu, S. Tolun, **A. Öztürk,** 2003. "Evolutionary Artificial Neural Networks in Hydrological Forecasting", *World Water and Environmental Resources Congress*, Philadelphia, USA.
- **A. Öztürk,** 2002. "Real Coded Genetic Algorithms in Optimization", MSc Thesis, Institute of Science and Technology, Istanbul Technical University.
- **A. Öztürk,** 1999. "Slope and Depth Perception Through Digital Imagery", BSc Thesis, Aeronautics and Astronautics Faculty, Istanbul Technical University.